Transmission Dynamics of Respiratory Syncytial Virus within the Household and in the Community

Thesis

UNRESTRICTED

Transmission Dynamics of Respiratory Syncytial Virus within the Household and in the Community

By

Emelda Aluoch Okiro, BSc

Sponsoring Establishment
KEMRI-Wellcome Trust Research Programme
Kilifi, Kenya

In collaboration with
University of Warwick

January 2007

Thesis submitted for the Degree of Doctor of Philosophy
Open University (UK)

Abstract

Introduction: Knowledge on RSV infection and disease is sparse in much of sub-Saharan Africa. Disease incidence is highest in infants, but there is little understanding on how and where they become infected. Characteristics of RSV transmission are poorly defined. Understanding transmission of RSV within the community is important in understanding infection and disease in infants and in defining potential effects of vaccination

Methods: A prospective longitudinal surveillance of ARI in a random selection of households in Kilifi District, coastal Kenya, was established in early 2003, and continued until 2005. Participants were under active and passive surveillance and were reviewed using a standard proforma. Nasal washings from symptomatic household members were screened for RSV antigen using IFAT. Oral-fluid for serological determination of infection was collected at enrolment and every 3 months.

Results: 81 households were recruited, 25 were lost to followup. 121 infections were identified approximately half of which were re-infections. The virus infected 54% of the households. Incidence of RSV infection was 218 cases/1000 cyo (95% CI, 182- 264). The incidence of primary infection was 476 cases/1000 cyo (95% CI, 361-630) and re-infection was 147/1000 cyo (95% CI, 115-189). Risk of disease was higher during primary infection than re-infection with the highest risk of disease in children 12-17months old. Estimated duration of viral shedding was short, <1 week. Age, sex, infection history and severity of infection were not found to significantly affect duration of shedding. Pre-school children with siblings in school had higher rates of infection. Index to secondary case interval was shorter for younger children and SARs were higher in smaller households. Crowding and stunting were associated with increased risk of both LRTI and RSV specific LRTI. Sanitation, type of house, and having siblings under the age of 6 years of age were associated with increased risk of RSV-disease.

Conclusion: We have defined several aspects of RSV transmission and demonstrated an important burden of RSV infection and disease in a rural Kenyan community addressing a need for more information from developing countries.

Acknowledgements

Firstly I would like to thank my two supervisors James Nokes and Graham Medley. James, for greatly inspiring me and for his constant encouragement. His knowledge of this subject was invaluable and his attention to detail has become ingrained. He has helped me appreciate the art of self criticism. Graham, for his remarkable insight and critical analysis and witty delivery of both. His encouragement made me believe in myself. I have learnt a tremendous amount from both and it has been a pleasure working with them. I could not have asked for better supervisors.

In Warwick- I would like to thank Lisa White who has taught me so much and provided tremendous support and advice through out this whole process. Lisa you and I both know how much you have helped me. I would also like to thank Dr Paul Scott for providing the ELISA plates and for his work on genotyping.

In Kilifi- I would like to thank the RSV team, the field workers Salma, Janet, Karen, Benson, Kadzo, Mwaruwa, Riziki, Milton, the Maternity Field workers, the Laboratory staff, John and Ann and the project Clinical Officer MJ all of whom worked tirelessly with enthusiasm throughout the whole project. Special thanks to Patience, the field worker responsible for the Household study. I would also like to thank Maureen and Marie (in Oxford) who helped with sourcing many of the papers I required. Many thanks to all the study participants and everyone at the Kemri Wellcome Trust Unit.

My friends: Emily, Linda, George, Beryl, Janet, Kendi, Joan, Oscar, Jane, Cleo, Edna, Sophie, Zena. Thank you for always being available to listen to me, to have a laugh and for offering encouragement. Rachel Ochola, for being both a friend and for going though this process together. I appreciate all your help with the ELISA work. Your support has been invaluable. All past students from Kilifi thank you for being excellent examples of success.

Most of all I would like to thank my family for their support, love and prayers. Mum and Dad, I have benefited greatly from your guidance, teaching and wisdom. Thank you for your unwavering support. Carol and Pat thank you for taking the time to read through parts of the thesis. My family has seen me through and I would not have got here without them.

Table of Contents

Chapter One 1

General Introduction 1

1.1 Defining the question 1

1.2 Thesis research objectives 4

1.3 Approach 5

1.4 Declaration of author's role 5

1.5 Overview of thesis 6

Chapter Two 8

Literature Review 8

2.1 Summary 8

2.2 Epidemiology of RSV 12

2.2.1 Seasonality 12

2.2.2 Characteristics of transmission 14

2.2.3 Molecular structure 16

2.2.4 RSV molecular epidemiology: Circulation of Groups A and B 18

2.3 Immunity to infection and disease and to subsequent re-infection 21

2.3.1 Viral characteristics that permit re-infection 23

2.4 Infection and re-infection 25

2.5. Severity of RSV infection in children 28

2.6 Re-infection 35

2.7 Household Studies 41

2.8 Risk factors for ARI and ALRI 45

2.9 Potential use of Oral Fluid to determine infection rates 56

2.10 RSV infection in Kenya 58

Chapter Three 61

Study Description 61

3.1 Overview 61

3.1.1 Chapter Aims 61

3.2 Study area 62

3.2.1 Geography 62

3.2.2 Epidemiological-Demographic surveillance system (Epi-DSS) 64

3.2.3 The Population....65
3.2.4 Climate....66
3.3 Methodology....68
3.3.1 Sample size calculation....69
3.3.2 Sampling design....69
3.3.3 Selection of study area....70
3.3.4 Surveillance of infection....71
3.3.5 Implementation....72
3.3.6 Surveillance methods....78
3.3.7 Study processes....80
3.3.8 Quality control (reliability)....83
3.3.9 Laboratory Methods....84
3.3.10 Data Handling....86
3.3.11 Ethical considerations....87
3.4 General Results....87
3.4.1 Recruitment....87
3.4.2 Loss to follow-up....88
3.4.3 Epidemics in the main birth cohort and in individuals of the household cohort 88
3.4.4 Surveillance intensity....90
3.4.5 RSV infection....92
3.4.6 Seasonality....92
3.5 Discussion....93
Chapter Four....103
Evaluation of the dynamics of RSV-specific antibodies in oral-fluid samples 103
4.1. Introduction....103
4.2 Chapter Aims....106
4.3. Study Design and Sample sizes....107
4.4. Oral-fluid Enzyme Linked Immunosorbent Assay (ELISA) and Optimisation....111
4.4.1 Optimisation....113
4.4 Discussion....126
Chapter Five....133

Transmission within the household..... 133

5.1 Introduction 133

5.2 Chapter Aims 135

5.3 Methods 136

5.4 Results 139

5.4.1 Infection and re-infection patterns 144

5.4.2 Introduction of infections into families 150

5.4.3 Spread of RSV infection within the home 152

5.5 Discussion 155

Duration of Viral Shedding 163

6.1 Introduction 163

6.2 Chapter aims 165

6.3 Methods 165

6.4 Results 171

6.5 Discussion 176

Chapter Seven 180

Age-stratified incidence of infection and disease severity 180

7.1 Introduction 180

7.2 Chapter Aims 181

7.3 Methods 181

7.4.1. Incidence of all respiratory illnesses 185

7.4.2 RSV infection and disease 189

7.5 Discussion 197

Chapter Eight 202

Risk Factor Survey 202

8.1 Introduction 202

8.2 Chapter Aims 203

8.3 Survey 203

8.3.1 Study Design and Ethical Considerations 203

8.3.2 Questionnaires 204

8.3.3 Data Management 205

8.3.4 Description of data set 205

8.3.5 Response rate 206
8.4 Analysis 206
8.4.1 Descriptive analysis 206
8.4.2 Creation of an asset based wealth index to estimate household socio-economic status 207
8.4.3 Calculation of anthropometric indicators based on WHO child growth standards.. 213
8.4.4 Regression models for cohort studies 214
8.4.5 Univariate Analysis 218
8.4.6 Multivariable Analysis 219
8.5 General Results 219
8.5.1 Lower Respiratory Tract Infection 223
8.5.2 RSV infection and RSV specific Lower respiratory tract infection (RSV-LRTI) 224
8.6 Discussion 224
8.6.1 Advantages and limitation or the study design and questionnaire 224
8.6.2 Validity of asset index 226
Chapter Nine 229
Determinants of Respiratory Disease at Household and Individual level. 229
9.1 Introduction 229
9.2 Chapter aims 229
9.3 Age-stratified incidence estimates 229
9.3.1 Interactions between age and gender, poverty level and HH density 231
9.4 Exploring data for RSVI, RSV-LRTI and all-cause LRTI outcomes 233
9.6 Summary of associations between variables 234
9.7 Analyses of Independent Effects 237
9.9 Goodness of fit of the Models. 249
9.10 Discussion 249
Chapter Ten 261
Overall Discussion 261
10.1 Introduction 261
10.2 Findings and evidence to support findings 262

10.3 Generalisability of study findings 267
10.4 Study Limitations 269
10.5 Recommendations for Further work.....270
References: 273
Appendix A: Exit Poll Questionnaire..... 300
Appendix B: Information and consent form for birth cohort child. 302
Appendix C: Information and consent form for household study participants..... 306
Appendix D: Information and consent form for Adults in the Household study..... 311
Appendix E: Information and consent form for RSV Study extension. ... 314
Appendix F: Active Surveillance Form..... 317
Appendix G: Field Work Procedures 320
Appendix G1: Flow Diagram 1 326
Appendix G2: Flow Diagram 2 327
Appendix G3: Flow Diagram 3 328
Appendix G4: Flow Diagram 4 329
Appendix G4: Flow Diagram 4 329
Appendix G5: Flow Diagram 5 330
Appendix G6: Flow Diagram 6 331
Appendix H: Sick child clinic visit questionnaire..... 336
Appendix I: Limitation on provision of services 341
Appendix J: Cytospin method for cell separation..... 343
Appendix K: The detection of Respiratory Syncytial Virus (RSV) using the direct immunofluorescence test. 344
Appendix L: RSV protocols for RT-PCR 349
Appendix M: Standard Operating Procedure for Oral-Fluid ELISA 355
Appendix N: SOP for Cell Culture and Preparation of Lysate..... 357
Appendix O: Calculating weights for observations within or between age classes by age strata 364

Appendix P: Incidence estimated by age and child class 366
Appendix Q: Incidence rates of RSV infection and disease stratified by age class 368
Appendix R: Effect of weighting on observations using different Age classifications 369
Appendix S: Risk Factors Survey Questionnaire 371
Appendix T: Guide for completing RSV risk factor study questionnaire 376
Appendix U: The means, ranges and distributions of variables used to compute the asset index 379
Appendix V: WHO Child Growth Standards 380
Appendix W: Results from univariate analysis 387
Appendix X: Testing the Proportional Hazard Assumption 401
Appendix Y: Assessment of Goodness of fit of the Cox regression models. 404

List of Tables

Table 2.1. Proportion of Respiratory Disease of Infancy and childhood Associated with RSV infection in various studies.....9

Table 2.2. RSV Proteins and their functions.....17

Table 2.3. RSV infection and re-infection rates in young children from studies in developed and developing countries.....26

Table 2.4. Adult challenge studies showing infection and illness rates after RSV infection 28

Table 2.5. Incidence rates of RSV-LRTI in developing countries in hospital studies.....31

Table 2.6. Incidence rates of RSV-LRTI in developing countries in community based studies 32

Table 2.7. Incidence rates of hospitalization for RSV among children < 1year in studies from Industrialized Countries 35

Table 2.8. Estimates of RSV associated re-infection rates.....37

Table 2.9. Summary of studies of defining respiratory syncytial virus infection within families.....43

Table 2.10. Factors associated with ARI and ALRTI: summary of findings form various studies.....47

Table 2.11. Factors associated with RSV Infection and disease: summary of findings form various studies 51

Table 2.12. Mean immunoglobulin concentrations (mg/L) in different salivary components 58

Table 3.1. The inclusion and any exclusion criteria for the households 74

Table 3.2. Reasons for exiting the HH study 88

Table 3.3. Details on surveillance intensity 91

Table 3.4. Details on nasal washings taken during the study.....92

Table 3.5. RSV Infection in household study 92

Table 3.6. Correlation between Fieldworker and Clinicians' clinical classification.....100

Table 4.1. Summary of the elements of the sampling involved in each study.....107

Table 4.2. Total numbers of samples collected for each of the 3 elements of serological study 111

Table 5.1. Table of definitions.....136

Table 5.2. Sex distribution of household individuals 143

Table 5.3. RSV infection in the household study by epidemic.....146
Table 5.4. Infections in siblings in the household cohort.....148
Table 5.5. Factors that influence outbreak size (number of infections in a household) in the household cohort.....149
Table 5.6. Characteristics of primary and secondary cases in RSV positive families with multiple infections by age, cohort type and school attendance.....151
Table 5.7 Relative frequency of introduction in households of different characteristics.....152
Table 5.8 Matrix of age of primary case and age of secondary cases 153
Table 5.9. Median interval between index and secondary cases of RSV infection in family children by age of the secondary case.....154
Table 5.10. Age-specific secondary attack rates according to the age of index cases, in episodes with a single index case 155
Table 6.1. Infectiousness period distribution (IPD) estimated using survival analysis using data from 192 infected children from the best fitting model (λ=10) 177
Table 6.2. Infectiousness period distribution estimated using survival analysis using data from 192 infected children with the more realistic value of λ=0.01 178
Table 7.1. Age and sex distribution of study participants at recruitment 188
Table 7.2 Incidence of respiratory illness according to age 189
Table 7.3. RSV cases by sex and age class including the weighted summaries of child years of observation 192
Table 7.4. Disease risks after primary and re-infection infection stratified by age 193
Table 7.5. IRR of RSV infection, RSV-LRTI and RSV-associated severe LRTI by age class 199
Table 7.6. Incidence estimate comparing weighted and un-weighted child years of observation 205
Table 8.1 Eigenvalues from the PCA. Showing the variance explained by the first 13 components 214
Table 8.2. First six components from the principal components analysis.....215
Table 8.3. Scoring factors for variables entering the computation of the first four principal components. 216
Table 8.4. Characteristics of population studied.....225
Table 8.5. Episodes of LRTI in cohort one and two 228

Table 8.6. Classification differences of the study population using different indices 231
Table 8.7 Comparison between the asset index and the profession of a household's major income provider 232
Table 9.1 Estimated hazard ratios for factors associated with clinical RSV infection 243
Table 9.2 Risk factors independently predicting increased hazard of RSV-LRTI / disease 245
Table 9.3. Risk factors independently predicting increased hazard of all-cause LRTI 247
Table 9.4 Summary table comparing individual risk factors across the different outcome groups is provided below.....250

List of Figures

Figure 2.1. Aetiology of ALRTI in developing countries in community based studies 10

Figure 2.2. Incidence of RSV-LRTI stratified by age for disease burden community studies 11

Figure 2.3. RSV IP cases by month in Kilifi, Jan 2002- December 2005 19

Figure 2.4. Analysis of genotypes from IP samples KDH 2002-2005.....20

Figure 2.5. Kilifi District Hospital inpatient surveillance of RSV.....60

Figure 3.1. Map of Kenya showing the location of Kilifi district 63

Figure 3.2. Comparative annual incidences of various diseases within the DSS over a five year period, 2002-05 by age group 65

Figure 3.3. Population pyramid of the Demographic surveillance system in Kilifi 66

Figure 3.4. Climatic conditions in Kilifi district between the years 2002-2005. Shows for each month the average daily precipitation (mm), relative humidity (RH), and maximum and minimum temperatures.....67

Figure 3.5. Schematic of the study design 68

Figure 3.6. Map showing spatial distribution of households within the DSS on a background showing the population density of the area from which the birth cohort was recruited 77

Figure 3.7. Nasal specimen collection by nasal wash (NW) bulb method in the clinic and in the field setting 81

Figure 3.8. Oral-fluid collection in the field using an Oracol 82

Figure 3.9. Temporal pattern of RSV infection in the birth cohorts. Cohort 1(a) 2(b). Grey area shows cohort size at each time point 89

Figure 3.10. The course of the epidemics in the HH study (blue bars) on a backdrop of recruitment and exit (grey area) from the study. IP cases are shown by the blue line.....90

Figure 3.11. Cases of infection in the Household cohort correlated to metrological data during follow-up 93

Figure 3.12. Frequency distribution of the second digits of respiratory rate variable.....101

Figure 4.1. Schematic of antibody dynamics (log scale) of RSV infection over a 3 month interval, with or without re-infection 106

Figure 4.2. Illustration of stages (1-4) of the IgG Capture ELISA assay 110

Figure 4.3. Illustration of staged 1-4 of the direct ELISA assay.....113

Figure 4.4. IgG standard curve 114

Figure 4.5. Absorbance readings for total IgG using doubling dilutions from 1/10 to 1/20480 for different oral-fluid samples. OD is optical density at 490nm 116

Figure 4.6. Absorbance readings from paired acute and convalescent oral fluid samples from Patient 56134.....118

Figure 4.7. Age plots for serum (left) at 1/100 to 1/400 dilution and oral-fluid (right) at 1/2 to 1/16 dilution.....119

Figure 4.8. Comparison of substrate processing times 10 vs. 30 minutes using samples from two adults.....120

Figure 4.9 Comparison of absorbance reading using different concentrations 5% (left) vs. 10% (right) of blocking solution.....121

Figure 4.10. Comparison of absorbance readings using two different secondary antibodies on both serum (right) and oral-fluid (left) samples.....122

Figure 4.11. Absorbance readings for paired serum (top) and oral-fluid (bottom) samples from a household child.....124

Figure 4.12. Oral-fluid (bottom panel) and serum (top panel) sample from Patient taken on four consecutive dates; 29-03-04, 27-04-04, 12-08-04 and 06-01-05 respectively.....125

Figure 4.13. Oral fluid checker board titration. Correct OD reading (Top panel). Bottom panel shows actual Ag and mock signals.....131

Figure 4.14. Evaluation of oral fluid for the detection of RSV infection: preliminary results..... 132

Figure 5.1. Transmission model of RSV.....134

Figure 5.2. Relationship between successive clinical cases 138

Figure 5.3. Frequency distributions of the interval from primary case onset to the onset of subsequent cases in households 139

Figure 5.4. Characteristics of households in the study at the beginning and at the end.....141

Figure 5.5a. The age distribution of siblings in the household study 142

Figure 5.5b. The age distribution of the adults in the household study 143

Figure 5.6 Distribution of individual infections (orange) and household infections (blue) by week 145

Figure 5.7. Infection (blue) and re-infection (maroon) patterns in relation to age in a household cohort.....147

Figure 5.8 Proportion of households into which virus introduced. Top panel shows number of households of each size.....150
Figure 6.1. Illustration of study design. Red arrows show sample collection 167
Figure 6.2. The time line for each study individual. The reference time is t_0 (day of first positive) which is set to zero for all individuals. t_{-1}, t_1 and t_2 are the times of last negative, last positive and first negative result respectively 168
Figure 6.3. The number of individuals at each time point in the shedding study 170
Figure 6.4. Probability of infection at different times assuming a truncated negative exponential distribution and infection occurs before 0. a) λ =0.01; b) λ =0.3 171
Figure 6.5. The Weibull probability distribution as the survivor function (blue) and the probability density function (black) (β= 3 and α=0.002) 172
Figure 6.6. Kaplan Meier survival estimates for covariates with highest significance.....175
Figure 6.7. Variation in results with the rate of infection (λ). a) the maximum log-likelihood; b) the parameters of the Weibull IPD function relative to their maximum value (solid line: α; dashed line: β); c) the mean (solid line) and standard deviation (dashed line) 176
Figure 7.1. Incidence (95% CI) of all-cause respiratory illness among children by age and sex 190
Figure 7.2. Incidence estimates of respiratory illness by age and child class 191
Figure 7.3. Proportion of primary infections resulting in disease (Risk) stratified by age class 194
Figure 7.4. Incidence rates of total infection (open bars), disease (light blue bars) and severe disease (dark blue bars) by age class. 196
Figure 7.5. Incidence rates of primary infection (open bars), disease (light blue bars) and severe disease (dark blue bars) by age class.....197
Figure 7.6. Incidence estimates of RSV infection by age and child class.....199
Figure 8.1. Follow up of cohort children by calendar time. Lines indicate individual follow-up of children. 220
Figure 8.2. Height for age of children in the study by sex and age. The *box plot* depicts the interquartile range as a *box* and the median as a line in the box. *Bars*, upper and lower adjacent values 226
Figure 8.3. Weight for age of children in the study by sex and age.....227
Figure 8.4. Weight for height of children in the study by sex and age 227

Figure 9.1. Age-stratified incidence of RSV infection (open bars), RSV-LRTI (light blue bars) and severe RSV-LRTI (dark blue bars) 235
Figure 9.2. Incidence of all-cause LRTI by age class with 95% CI 236
Figure 9.3 Age-stratified incidences of RSV-LRTI, RSVI and LRTI cases by poverty classification: poor, middle and rich 237
Figure 9.4. Age-stratified incidence of RSV-LRTI, RSVI and LRTI by household density 238
Figure 9.5. Correlation between family density and number of children in the household (y-axis) for households of difference socioeconomic class.....242
Figure 9.6 Plot for predictor number of children in the home including the reference line at y=0 253
Figure 9.7 Plot for predictor number of siblings 6+ years old sleeping in same room as index_2 including the reference line at y=0 253
Figure 9.8 Association between malnutrition indicators and SES. X-axis denotes socio-economic classification - 1-poor, 2 - middle and 3 - rich class 256

Abbreviations

Ag / Ab	Antigen / Antibody
ARD	Acute respiratory disease
ARI	Acute respiratory infection
ALRI	Acute lower respiratory infection
bcc	Birth cohort child
BOSTID	Board of Science & Technology in Development (Unit of the US National Academy of Sciences)
cs	Cox-Snell
cy	Child years
cyo	Child years of observation
CL	Confidence Limits
DSS	Demographic surveillance system
ELISA	Enzyme-linked immunosorbent assay
FW	Field worker
haz	Height-for-age z-score
HH	Households
IFAT	Direct Immunofluorescent Antibody Test
IgG	Immunoglobulin G
IgM	Immunoglobulin M
IP	In-patient
IR	Incidence rate
IRR	Incidence rate ratio
KDH	Kilifi District Hospital
KM	Kaplan Meier
LRTD	Lower respiratory tract disease
LRTI	Lower respiratory tract infection
MCH	Maternal Child Health
MIP	Major income provider
NPA	Nasopharyngeal aspirate
NW	Nasopharyngeal/ Nasal wash
OD	Optical density

NW	Nasopharyngeal/ Nasal wash
OD	Optical density
OP	Out-patient
OR	Odds ratio
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PCT	Primary care taker
pfu	Plaque forming units
RF	Risk factor
RFLP	Restriction Fragment Length Polymorphism
RT PCR	Reverse Transcriptase PCR
RSV	Respiratory syncytial virus
RSVI	RSV Infection
RSV-LRTI	RSV specific lower respiratory tract infection
RSV-s-LRTI	Severe RSV specific lower respiratory tract infection
sLRTI	Severe LRTI
SAR	Secondary attack rate
SES	Socio-economic status
SOP	Standard Operating Procedure
URTI	Upper respiratory tract infection
WAIFW	Who aquires infection from whom
WHO	World Health Organisation

Chapter One

General Introduction

1.1 Defining the question

Respiratory syncytial virus (RSV) is the major viral cause of severe childhood acute respiratory infections (ARI) worldwide. Most disease occurs in infants and young children making them an important group that would be a target population for vaccines. The last few years have seen significant progress in the development of a vaccine against RSV [1-3]. A recent success is that of a recombinant live attenuated vaccine which has been shown to be sufficiently attenuated and apparently still immunogenic in the target age group, infants 1-2 months old [2]. This vaccine will potentially play a major role in the control of disease rather than providing sterilizing immunity to infection (it is unlikely to elicit herd immunity in a locality.)

To correctly evaluate the potential effect of a vaccine in a population, it is necessary to understand the infection dynamics prior to vaccination. However the characteristics of RSV transmission are poorly understood - its persistence and recurrence, re-infection, potentially predisposing risk factors, age-related risks, role of household contacts, and the effect of prior exposure and past strain infection are not well defined. Much of this information is lacking in developed countries and is in particular shortage – practically absent – from developing countries.

Previous studies from developing countries [4-10] report widely varying incidence of RSV-associated lower respiratory infection (RSV-LRI) between 8/1000 cyo to 220/1000 cyo. While population or biological factors might account for some of this variation, methodological issues provide a more likely explanation. More recently, 4 studies arising

from the same generic protocol (although not strictly adhered to) were undertaken [11]. From these studies the incidence of RSV-LRI in children < 5 years was 34/1000 and 94/1000 in Indonesia and Nigeria, respectively. The incidence of severe RSV-LRI per 1000 child years was 5 in Mozambique, 10 in Indonesia and 9 in South Africa with the majority of RSV cases occurring in infants. Though generally lower, two studies in the developed world (one in low income families in the US [12]) have reported similar rates [13]. Although variable, these results make a substantial case for the burden of RSV-LRI as well as creating a need for the investigation of putative individual and household risk factors which may alter a child's risk of infection and disease. This is of general interest as children bear the greatest burden of disease and should be the main beneficiaries of a vaccine.

It appears unlikely that primary infections, while resulting in most disease [12], drive transmission; infants mostly stay in the home, and it's more likely that they are in fact the recipients and not the suppliers of infection. Transmission of RSV is known to be via droplet transmission and fomites [14] implying that close contact is necessary and that family transmission may play an important role in community transmission. Longitudinal studies of households can thus serve to describe aspects of the spread and transmission of infection. Our interest in older siblings and other household members (householders) is because they are probably the source of infection generally, and specifically the source for infants.

Household studies [12, 15-17] and community studies [18-20] from developed countries show RSV to re-infect repeatedly throughout life which points to a potentially significant role for re-infections within the population that may be fundamental to RSV persistence within the community. Finally, RSV is typical of many viral infections with respect to inducing less than solid immunity, and being antigenically and genotypically diverse. Hence the general

questions of persistence and the importance of re-infection in the transmission process are important to ask.

The research presented in this thesis was part of and nested within the main RSV birth cohort project. The aim of the overall project was to undertake longitudinal epidemiological studies within a well defined rural Kenyan population to further the understanding of RSV transmission dynamics and immunity. In particular the objectives were to:

1. Quantify through active and passive surveillance of an infant cohort the protective efficacy to re-infection of past history of infection, in particular homologous or heterologous strain infection and also a) single or multiple infections b) time since and, c) age at, infection.
2. Describe differences between RSV infections identified through hospital surveillance and those through community surveillance, with respect to spatial, temporal and genotypic characteristics.
3. Estimate rates of RSV infection and re-infection in the community in all ages, and elucidate the seasonal incidence characteristics of the virus (i.e. where is the virus circulating between major seasonal epidemics).
4. Advance understanding of the antibody responses to the immunologically important attachment (G) protein in RSV variants, in relation to infection and re-infection.
5. Elucidate, using results from the above combined with mathematical modeling, the mechanisms underlying the observation on spatio-temporal dynamics in RSV genotypes.

The research reported in this thesis was designed to contribute to the overall project by generating key information on the community dynamics of RSV infection by the

identification and quantification of who is being infected, how much infection (and re-infection) is occurring in the different age groups, how they contract infection, who acquires infection from whom (WAIFW), and how severe is the resultant disease. These data will be established in a context of the relationship to past exposure, infecting group (antigenic 'strain') and genotype and level of shedding from the infected individual.

1.2 Thesis research objectives

The general objective was to develop understanding of the mechanism of RSV transmission and related influencing factors, within a Kenyan population, in a spatial-temporal setting and with emphasis on events within the household.

Specific objectives *(a) Within households*

1. Describe RSV re-infection patterns in relation to age, family structure, RSV group and genotype, season, and environmental risk factors.
2. Monitor the time course of sequential infections to shed light on the source of primary RSV infections in infancy.
3. Estimate the duration of shedding of RSV following infection as an important transmission factor, in relation to influences such as age and previous infection.
4. Quantify incidence rates of RSV infection with respect to age.
5. Identify the occurrence of transmission outside of the main epidemic season to elucidate the mechanism of RSV persistence.
6. Investigate severity of RSV infection within the household in relation to age, virus genotype and household structure.

(b) Within a birth cohort

7. Identify and compare household and individual risk-factors for all cause LRTI and RSV infection and disease.

1.3 Approach

In the present research there was a detailed surveillance of infection in a birth cohort which was used to identify risk factors for clinical RSV infection and RSV-LRTI as well as those for all-cause lower respiratory tract infection. This cohort study provided an ideal opportunity to explore transmission at the household level by recruiting households of a sub-sample of the main cohort study.

1.4 Declaration of author's role

The study was designed by the principal investigators; Dr James D. Nokes, Prof. Graham Medley and Dr Patricia Cane. The author of this thesis was the field coordinator responsible for the overall management of the projects field activities, including; overseeing recruitment of study participants, follow-up, data and specimen collection. The author was also responsible for creating the risk factor database within the main study database. The household study was proposed in the original project but without great detail and the author was co-responsible for the design, and took lead in the study implementation. The author managed a team of 9 field workers and was responsible for managing the day-to-day activities and for task allocation. Unless otherwise stated the author took lead in the analyses presented in this thesis. Publications that have so far arisen from the project are listed below:

1. Nokes, D. J., **Okiro, E. A.**, Ngama, M., White, L. J., Ochola, R., Scott, P. D., Cane, P. A. and Medley, G. F. Respiratory syncytial virus epidemiology in a birth cohort from

Kilifi district, Kenya: infection during the first year of life. *J Infect Dis*, 2004; 190: 1828-32.

2. Scott, P. D., Ochola, R., Ngama, M., **Okiro, E. A.**, Nokes, D. J., Medley, G. F. and Cane, P. A. Molecular epidemiology of respiratory syncytial virus in Kilifi district, Kenya. *J Med Virol*, 2004; 74:344-354.

3. Scott, P. D., Ochola, R., Ngama, M., **Okiro, E. A.**, Nokes, D. J., Medley, G. F. and Cane, P. A. Molecular analysis of respiratory syncytial virus reinfections in infants from coastal Kenya *J Infect Dis*, 2006;193: 59-67.

4. **Okiro, E. A.**, Ngama, M., White, L. J., Cane, P. A., Medley, G. F. and Nokes, D. J Estimation of the duration of shedding of Respiratory syncytial virus (In Preparation)

1.5 Overview of thesis

A review chapter follows directly after the Introduction. The aim of the review is to provide an introduction to RSV, its epidemiology and transmission by which to establish the background from which the questions in this thesis are raised. An overview of putative risk factors for RSV and those for all-cause lower respiratory infcetion (LRTI) is also presented. Chapter 3 describes the overall study design giving specific details on methods of surveillance and all study processes both field and laboratory. Some general results on the description of households and study population, the completeness of follow-up (dropout rate), surveillance intensity and occurrence of epidemics are also presented. The fourth chapter presents an evaluation of a non-invasive method for examining rates of infection, through the detection of changes in levels of RSV-specific IgG in oral-fluid samples. The fifth chapter provides general results from the household surveillance study on patterns of infection and correlates of transmission within the household. Secondary attack rates and the relative risk of introduction of infection into the home by different age categories of individuals are derived

and discussed. In Chapter 6 the estimation of the duration of shedding using a survival analysis approach is described. Chapter 7 details the age-stratified incidence estimates for this study population, based upon clinically apparent cases confirmed by antigen assay of nasal washings. This analysis does not include evidence from non-invasive methods, because the identification of infection through oral-fluid samples proved unsuccessful. In Chapter 8, details of the risk factor survey and a background of the statistical methods used to analyse this data are presented. In the ninth chapter results of the analysis of the risk factor data are presented. This data is used to determine the risk factors for RSV infection (RSVI) and for RSV specific lower respiratory tract infection (RSV-LRTI) which is then compared with those risk factors for all cause LRTI. The final chapter provides a summary and discussion of the main finding of this thesis, considers possible improvements on the study and discusses the direction of future research.

Chapter Two

Literature Review

2.1 Summary

Acute respiratory disease (ARD) in infants is a worldwide public health problem [21, 22].

The World Health Organization (WHO) estimates the annual deaths in children under 5 years to be 12.2.million, one third of which are attributed to acute infections of the lower respiratory tract (ALRIs - pneumonia, bronchiolitis and bronchitis) [23]. Other studies have since confirmed that ARIs cause a considerable percentage of childhood mortality [24] with the highest percentage occurring in Africa. More recently, Bryce *et al* [25] reported that 19% of childhood deaths worldwide are caused by pneumonia. In Kenya, ARIs are the second leading cause of morbidity (after malaria) accounting for up to one quarter of outpatient attendance in health facilities, thereby being a key health concern [26]. Causative organisms for LRTI may be bacterial 4.5%- 40% (most commonly *Streptococcus pneumoniae* and *Haemophilus influenzae*) or viral 14% -64% [27].

RSV has been described as the single most important viral cause of ALRI in children particularly in those younger than six months [27-29]. In addition, RSV infections are responsible for most cases of severe symptoms such as bronchiolitis and pneumonia and for high numbers of hospitalisations, as for example shown in classic studies in the US (Table 2.1.)

Table 2.1. Proportion of respiratory disease of infancy and childhood associated with RSV infection in various studies

Diagnosis	*Study, Location*	*Virus recovery*		*Serological Evidence of Infection*	
		No. tested	% positive	No. tested	% positive
Brochiolitis	NIH study [30] Washington D.C	154	25	230	18
	[31] Chicago	34	50	-	-
	[32] Philadelphia	-	-	26	38
	[33] Columbia	48	42	39	67
	[34] Rhone Alpes	-	45	-	
	[35] Washington D.C	-	43	-	30
	[36] India	-	58	-	-
	[37] Tuscon	-	65	-	-
Pneumonia	NIH study [30] Washington D.C	278	10	522	14
	[31] Chicago	22	36	-	-
	[32]Philadelphia		24		39
	[33] Columbia	58	-	45	44
	[35] Washington D.C	-	25	-	22
	[36] India	-	19	-	-
	[37] Tuscon	-	33	-	-
Other acute respiratory diseases	NIH study [30]Washington D.C	1999	4.4	678	10
	[31] Chicago	-	15	-	
	[32] Philadelphia	-	-	481	16
	[35] Washington D.C	-	9	-	18

Most of the information that is available on RSV is from studies of paediatric patients admitted to hospital with serious respiratory disease and these have formed the basis for understanding the behaviour of RSV in young children. These data are useful but have limitations which include: i) the assumption that all serious cases get to hospital; an issue that is of particular concern in developing countries where a minority of cases actually make it to hospital [5, 9] and ii) ignores the individually less serious, but still important, disease burden in the community in general. Relatively few community based studies of RSV have been carried out [12, 15, 16, 19, 37], and even fewer still have been set in the developing world (Figure 2.1).

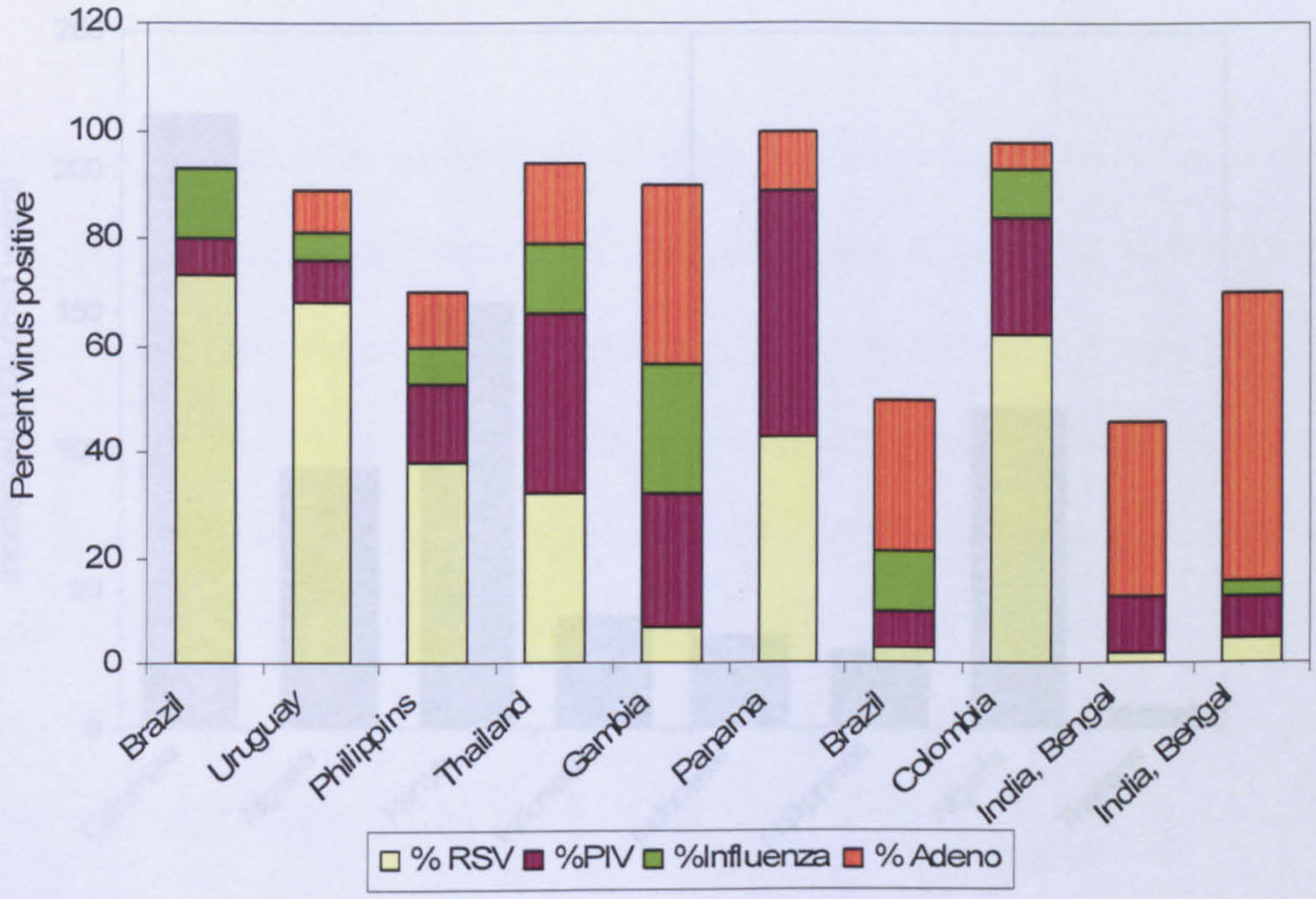


Figure 2.1. Aetiology of ALRTI in developing countries in community based studies (adapted from review by Weber *et al* [28]). Note: 5 studies on right hand side did not use immunofluorescent antigen test (IFAT); use of IFAT increased proportion of RSV by twofold in the other studies. PIV- Parainfluenza virus, Adeno- Adenovirus

Morbidity attributable to RSV in the community is not negligible as evidenced by denominator based disease burden estimates from community studies in developing countries summarized in Figure 2.2 and discussed further in section 2.5.

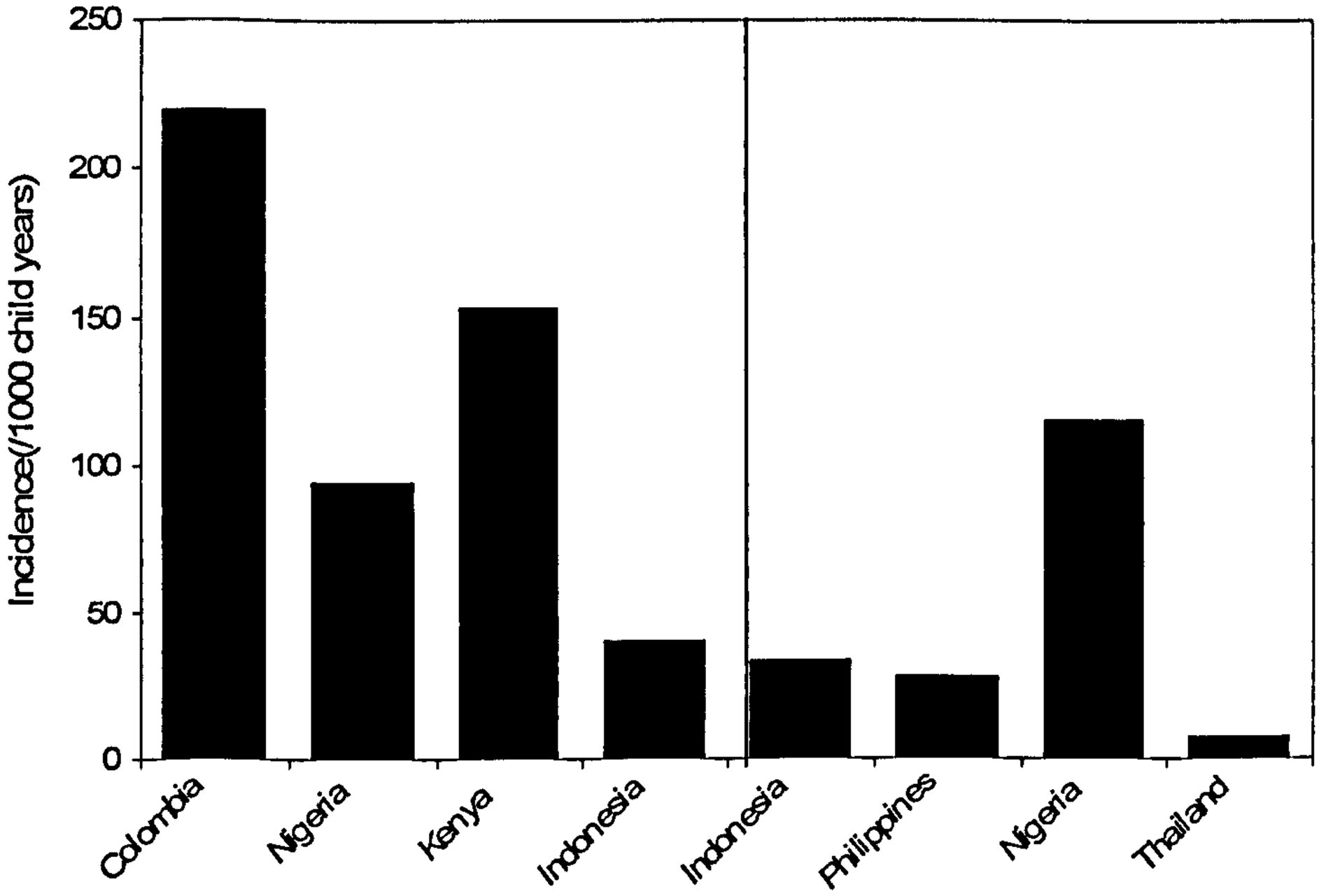


Figure 2.2. Incidence of RSV-LRTI stratified by age for disease burden community studies (adapted from review by Nokes [38])

Case fatality due to RSV is deemed generally to be low but has been shown to be higher in developing countries [28, 39]. Two reviews of studies in tropical and developing countries [28, 39] reported mortality rates of between 0-12% with a mode of 0%. However, estimated magnitude may be in question; in one Indonesian study [8] the majority of deaths of children admitted with pneumonia did so before a viral diagnosis could be undertaken; this potential bias may not be uncommon in resource poor settings. Furthermore, RSV causes considerable burden on health resources accounting for up to 40% of ARI admissions during RSV epidemics (unpublished data from Kilifi, DJ Nokes). This has significant implications to availability of beds and use of resources (such as oxygen for supportive therapy) that are scarce in developing countries.

Currently there is no vaccine in use although a recent recombinant live attenuated vaccine has demonstrated sufficient attenuation while retaining immunogenicity in the key target

age group (1-2 months) for the first time [2]. It is possible that in the not too distant future, later phase clinical trials will be undertaken and these will invariably benefit from improved data on disease burden estimates, re-infection patterns and community transmission all of which are at present scarce in the developing countries. Data from this study will go towards improving the knowledge on RSV epidemiology in developing countries.

2.2 Epidemiology of RSV

2.2.1 Seasonality

The epidemiology of RSV infection and disease is characterised by marked seasonal patterns. It is not entirely clear what triggers these epidemics or what happens to the virus in the inter-epidemic period but it is apparent that the timing of epidemics varies considerably. Patterns of transmission and persistence differ strikingly; ranging from typically long (~6 months) annual epidemics followed by periods of fade out to shorter discrete epidemics of one month [40], to reported cases of year round occurrence [41]. Repeat infection with RSV is common throughout life despite previous exposure which suggests that transmission should not have this strong seasonal pattern. This contribution of re-infection to the circulation of the virus, both during and between seasonal epidemics (of infection and disease), within the community has not been well defined.

Two reviews have documented the seasonal patterns of RSV [28, 39]. In general, epidemics occur annually during winter in temperate countries with the exception of some countries e.g. Finland where every two years there is a minor peak in April followed by a major peak in December [28, 42-44]. In areas with tropical or subtropical climate and seasonal rainfall, RSV outbreaks were more frequently associated with the rainy rather than the colder season

[28]. Countries closer to the equator with high rainfall through out the year were found to have a different seasonal pattern with most cases appearing in one half of the year but not in the other. In the review by Stensballe [39], seasonality in the Northern tropical areas was associated with a decrease in temperature and an increase in rainfall [45, 46] while epidemics south of the equator were associated with both decreased temperature and rainfall [47, 48]. Epidemics can occur completely out of phase in areas not far apart and with similar weather patterns [11, 28]. In Kilifi, unpublished surveillance data collected over the last five years indicates that RSV seasonality does not appear to be defined by the rainfall patterns (Chapter 3) but seems to exhibit a pattern similar to that seen in Finland.

From these findings it appears that neither temperature nor rainfall are main determinants of the timing of these outbreaks. What seems clear is that the seasonal pattern of RSV is multifaceted without a clearly defined mechanism. Climatic and geographic factors appear to play a part although not exclusively, pointing to the possible role of other social-behavioral or viral characteristics as factors that may contribute to this epidemic pattern. Socio-behavioral factors, for instance less indoor crowding or vacationing of school children, may play a part [49]. Several studies done in the developed countries provide evidence that younger children acquire infection from school-aged children within the household [15, 16, 50], indicating a significant contribution of within-school transmission to the dynamics of the epidemics. Waris and White [49] indeed make such a case pushing forward the idea of the school year as a possible explanation for the seasonal signal of RSV infection in many countries. This idea has at present not been explored in a developing country study and forms an aspect of one of the objectives of the planned study.

This seasonal pattern of RSV is also likely to be affected and /or determined by its natural dynamics. This includes the combination of the short-lived immunity to RSV infection resulting in a build up of susceptibility to re-infection (explored further in section 2.3) and new births creating cohorts also losing maternal antibody protection. These concepts and other possible seasonal triggers lead into the molecular epidemiology and antigenic variation exhibited by RSV, and how these may influence its epidemiological features (this aspect will be covered in further detail in section 2.2.4). Its mode of transmission (discussed below) doesn't seem to correlate with its rapid spread within communities during epidemic. RSV requires closes contact for spread to occur and is also shed for relatively short durations which would imply restricted spread of infection.

2.2.2 Characteristics of transmission

The major mode of spread is by large droplets or through fomite contamination [51]. Aerosol particles are not considered to be a major mode of spread, since the virus is not stable when aerosolized [52, 53]. Spread thus requires close contact with infected people or contamination of the hands with fomites from surfaces to which infected respiratory secretions have spread [51, 54]. Past studies clearly demonstrate that RSV is efficiently transmitted within hospitals apparently due to this propensity for transmission on fomites [14, 55, 56]. This then implies that social contacts are also important for spread. RSV is unstable in the environment surviving only a few hours on environmental surfaces, and is readily inactivated with soap, water and disinfectants [54].

Experimental challenge studies and duration of shedding studies report that most people shed virus for between 1-3 days before clinical symptoms with a great majority of illness onsets following initial viral recovery [57-59]. Individuals developing symptoms were

observed to start shedding virus earlier than those asymptomatic; 3-4 days vs. 5+ days post inoculation. Most children recover from illness in 7 to 15 days [51, 54]. This duration of infection and its variability is an important element in the dynamics of infection transmission. Hence, it is important to quantify the infection / recovery process, and in particular to know how past infection modifies the infectiousness of individuals [60].

Published estimates of the duration of shedding of RSV are few. A number of studies done were from children hospitalized with acute respiratory infection (ARI) [61-64]. Such estimates tend to be biased towards infections causing severe disease (arising from ~1% of primary cases) and from a narrow (young) age group. It has been shown that greater severity of infection within in-patients results in increased duration of shedding [62]. A generalization of these hospital-based estimates to all infections occurring within a community is unlikely to be accurate. One study estimated the duration of RSV shedding through home based monitoring, a study of individuals of all ages in 36 families in the United States during a single epidemic [15]. A mean duration of 3.5-7.4 days in otherwise healthy children was recorded (range 1-36 days), with higher duration for children under 2 years than for children under 16 years (9 vs. 3.9 days). Estimates, however, were not defined in relation to past infection status which is an aspect we intend to explore in the present study. A second family study following children under 4 years of age reported over 70% of cultures positive up to 7 days post illness onset, falling to less than 10% in week 2 and 3 [59]. No evidence for a difference in shedding duration between primary and re-infections was observed in this study (although with only 12 re-infections the power to detect even quite a large difference would be small). Studies of adults [65] in the community show shedding to be of similar or slightly shorter duration than shedding for children 1.6-3.9 days (range 1-27).

Despite the fact that the median duration of RSV shedding tends to be of the order of days, considerable variation exists: in the family studies referred to earlier, 5-10% of individuals shed for more than 14 days, and periods of shedding of 30-40 days have been recorded [15, 59, 62, 63]. Similarly, immunocompromised individuals have been known to shed RSV for longer possibly keeping RSV ticking over in low season [41]. Given the prevalence of HIV infection in much of sub-Saharan Africa, this issue of long term shedding needs to be investigated and quantified as it has implications epidemiologically for persistence [39]. This led us to investigate the duration of shedding in individuals within household in relation to past exposure – number of previous infections, severity, age and infecting group (Chapter 6).

2.2.3 Molecular structure

RSV is antigenically diverse with group and genotype structure an aspect that potentially affects its epidemiolgy and is explored further in the next section. There are two major antigenic categories of RSV, known as group A and B. Group B viruses are less variable than the group A viruses [66, 67]. The RSV genome contains 15200 nucleotides that are transcribed into 11 major subgenomic mRNAs each coding for a protein (Table 2.2).

Table 2.2. RSV Proteins and their functions (modified from Hacking *et al* [68])

Protein	Function
Non-structural proteins	
NS1/2	anti-interferon α and β activity
Nucleocapside proteins	
N	Nucleoprotein essential for transcriptional activity
P	Phosphoprotein essential for transcriptional activity
L	RNA polymerase
Transmembrane glycoproteins (surface proteins and important in immune stimulation)	
SH	Small Hydrophobic protein: function unknown
G	Glycoprotein: viral attachment to cell
F	Fusion protein: viral entry and syncytia formation
Matrix proteins	
M	Matrix protein: viral assembly
M2	M2-1: transcription elongation factor
	M2-2: regulation of viral transcription

The G and F proteins are the major antigenic determinants of the virus because they stimulate the production of protective immune response. Both these proteins can induce the production of potent RSV-neutralizing antibodies by host cells [69] but only the F stimulates signifcant cytotoxic T-lymphocyte responses [55, 70]. It is less clear what the relative roles of F and G are to the strain structure observed as the virus has been shown to replicate and reinfect (in animal studies) in the absence of the G protein [71]. Still, it is the G protein that seems to be variable under immune pressure due to the high proportion of nucleotide changes that result in amino acid coding changes [72, 73]. The major antigenic and nucleotide sequence differences between the two groups are thus found on the G attachment glycoprotein [55, 74]. There is further genotypic structure, with numerous

genotypes within each group [75] but the functional importance of this structure is not known. It is thought that this antigenic variation both within and between groups may enable re-infections to occur [76] (explored further in section 2.3.1). However, it is unclear what cross-immunity exists between the two Groups - it is known that the subtypes have antigenic cross reactivity for the F protein [77]. This is an aspect that is explored within the main project (Paper by P. Scott listed above). Molecular epidemiological studies (discusssed below) show evidence for interaction between Groups and genotypes i.e. clear patterns of A and B prevalence and sequential replacement of genotype variants in some longitudinal studies.

2.2.4 RSV molecular epidemiology: Circulation of Groups A and B

The epidemiological picture is complicated by the antigenic group variation [78-81]. RSV groups A and B co-circulate; however, there is usually a distinct temporal pattern of frequency. For example, long term surveillance of hospital admissions in Birmingham UK show two years of higher prevalence of RSV A followed by one year of RSV B [42, 75]. In Finland another example, the pattern is of two epidemics of one type followed by two of the alternative group [44]. In general, years in which Group A dominates (i.e. is of higher frequency) exceed those for Group B [66, 67, 81-83]. The pattern observed in Finland is more complex insofar as 2 epidemics (one small and one large) of the same Group occur in close succession (average interval of one month or less) followed by a longer inter-epidemic period (~7 months). The same pattern has been seen elsewhere in Scandinavia and North European countries [84], and a similar pattern appears to be observed in Kilifi (Figure 2.3 & 2.4).

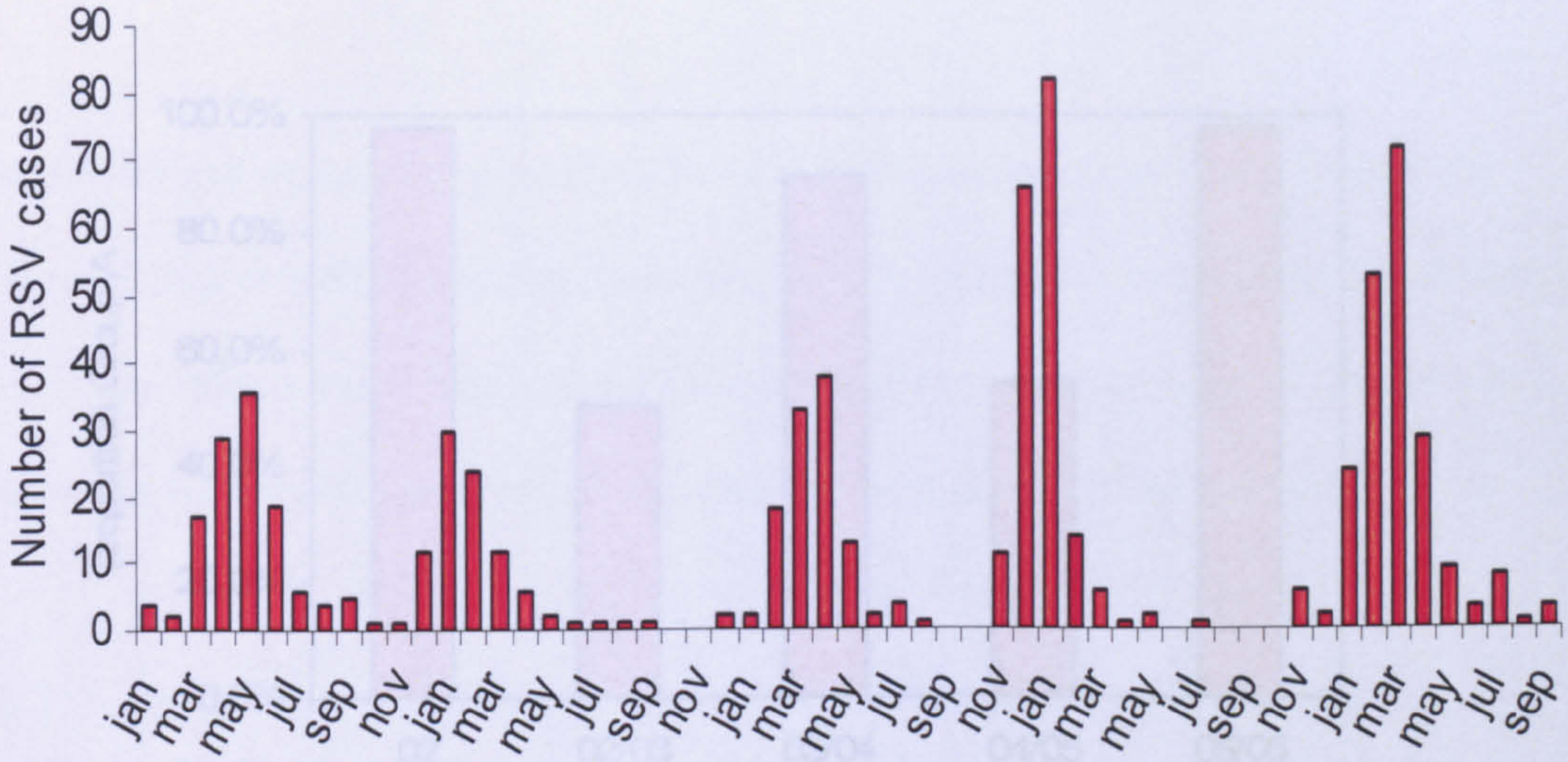


Figure 2.3. RSV IP cases by month in Kilifi, January 2002- December 2005

This sequential dominance suggests that there is an underlying non-random structure at play; possibly some level of immunity forcing these shifts. Cane *et al* [42] hypothesize that strain differences affect protective immunity. In this case a novel virus may be transmitted more efficiently thus resulting in a situation where the epidemic strain would be determined by the levels of strain specific immunity in the community, with different social structures explaining the different patterns observed between communities.

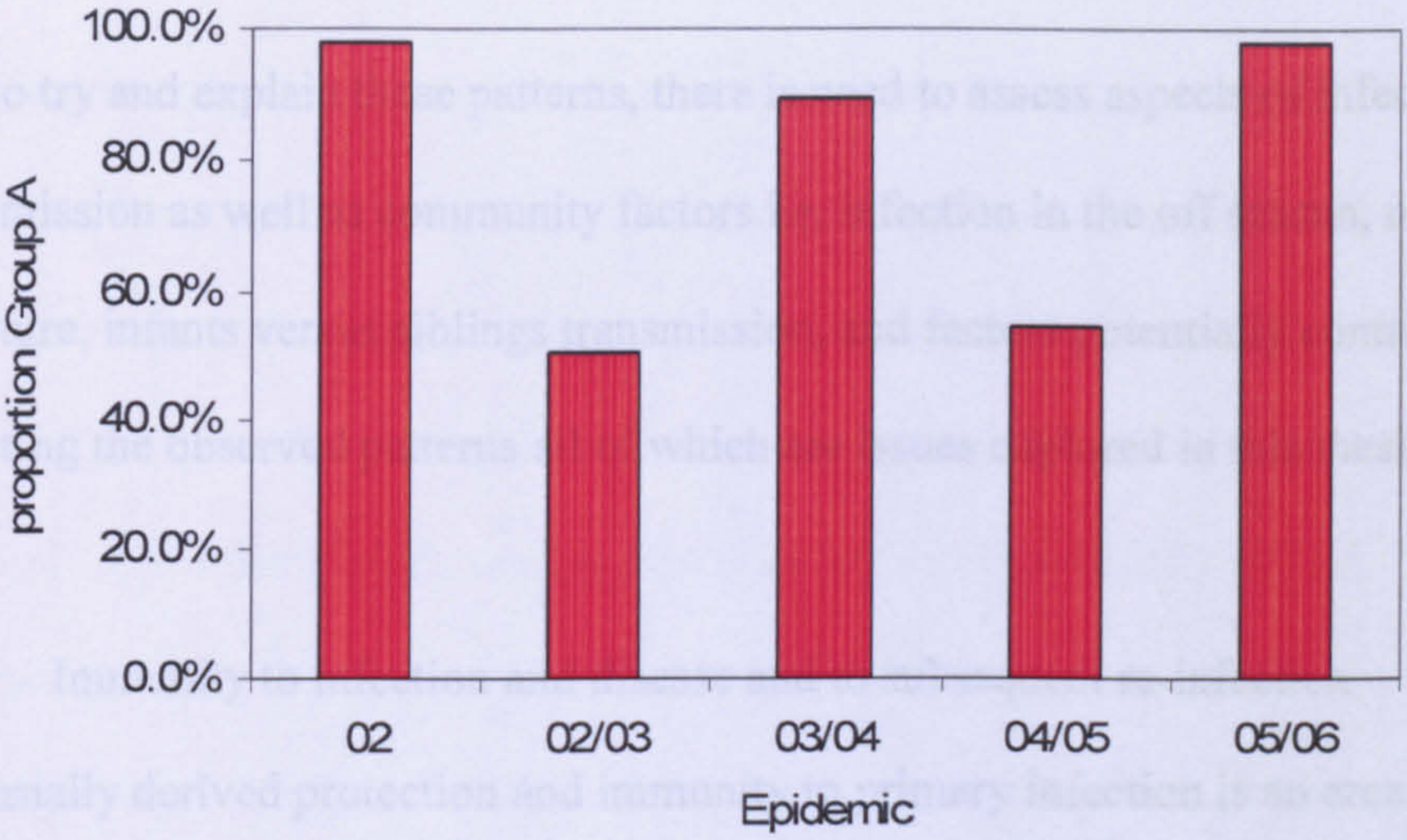


Figure 2.4. Analysis of genotypes from IP samples KDH 2002-2006. >85% of the samples were successfully typed. Group A viruses are dominant in alternating years.

The consistency of the patterns is indeed suggestive of an underlying mechanism. Modelling approaches have been used in an attempt to better understand the processes at work. Mathematical models capturing the host-pathogen interaction and the population dynamics (immunity) were explored by White *et al* and Weber *et al* [60, 85]. The work of White *et al* [60], suggests the structure to be the result of (i) A being more transmissible than B and (ii) differential within to between group specific immunity. Immunity was estimated to be greater for homologous challenge (60%) than heterologous (16%). Seasonality in transmission is known to relate to contact rate for measles [86, 87] and was in this case dependent on waning immunity. From mathematical models, it appears that several different assumptions can explain the observed patterns thus providing no definitive clarification on this seasonal phenomenon. These observations together would suggest a potentially complex relationship between the viral genetic variation, strain specific immunity and how these relate to the process of infection / disease as well as to transmission dynamics. In order to further understanding on RSV transmission dynamics

and to try and explain these patterns, there is need to assess aspects of infection and transmission as well as community factors i.e. infection in the off season, role of family structure, infants versus siblings transmission, and factors potentially contributing to and/or affecting the observed patterns all of which are issues explored in this thesis.

2.3 Immunity to infection and disease and to subsequent re-infection

Maternally derived protection and immunity to primary infection is an area that remains controversial as there are contradictory results about the role of maternal antibody in protection. Antibody found in serum of newborn infants represents passively acquired antibody, and commonly last less than six months [88, 89]. The high incidence of infection and disease in infants appear to support the hypothesis that maternal antibodies do not offer complete protection. In fact it was previously suggested that maternal antibodies interfered with the immune response of the infants resulting in more severe disease [90] similar to what was seen with the formalin vaccine where vaccinated infants had more severe natural infection [91]. Subsequent studies since have given differing results. In some cases, the severity of illness resulting from RSV infection in the first months of life has been shown to be modified by high levels of maternal antibody [92-94], while in other studies primary infection in infants involving the lower respiratory tract has been shown to occur despite the existence of passively acquired maternal IgG antibodies in the circulation [95]. More recently, prophylactic administration of RSV-specific immunoglobulin (pooled immunoglobulin with high titers of anti-RSV antibody) given to high-risk infants and young children has been shown to be effective in reducing severe disease (by reducing the number and duration of hospitalizations) due to RSV [96, 97] arguing strongly in favour of the protective effect of RSV antibody. Rather counter intuitively, this RSV-specific immunoglobulin provides no protection when administered during acute infection [52].

RSV is primarily a mucosal virus thus immunoglobulin A (IgA) is also likely to be important for protection. A study of RSV-specific IgA in community dwelling elderly persons, those with congestive heart failure and chronic pulmonary disease, and healthy young adults (19-40 years) with respiratory illness found that low RSV specific nasal IgA was an independent significant risk factor for RSV infection [98]. Similar findings were reported from the studies by Mills *et al* [99] and Watt *et al* [100] which involved the experimental challenge of healthy adults. Mills *et al* found that high titers of nasal antibody were correlated with lower titres of virus shed and with resistance to illness. Nasal neutralizing antibody was also found to be a better correlate for protection than serum neutralizing antibody. In yet another study [101], the presence of IgA was to some extent correlated with resistance to re-infection though this finding was not consistent in all instances. Seemingly from these findings IgA levels are an important protective factor.

As is the case with maternal acquired antibody, the presence of antibodies in individuals who have had a previous infection in general does not seem to be protective against infection, although the risk of re-infection has been associated with the number of previous infections and level of pre-existing antibody [12, 20, 101]. Systemic neutralizing antibodies (to F and G) are produced in increasing concentrations in response to annual infections of the same or different strains in young children [102] with serologic surveys showing that levels of neutralizing antibodies increase with age [89, 103]. Several studies show an association between increasing levels of humoral antibody and increasing protection from re-infection [12, 20, 101] and/or disease following re-infection [12, 44]. Two of these studies used microneutralization assays on serum samples while the other used an RSV specific IgG ELISA on saliva samples [20]. The study by Hall *et al* was an adult challenge study with 15 subjects repeatedly challenged. The other studies were of children followed

up from birth [12] and school children [20] both with larger sample sizes; 123 and 121 children respectively. These data in addition to the observation that infection rates decrease with age [44] provide evidence that serum antibody appears to offer some protection against disease and or infection albeit never completely. One study [104] was able to define a minimum protective threshold against RSV-associated-hospitalization. Evidence from this study suggests that neutralizing antibodies of $\geq$ 6.0 $\log_2$ (arithmetic titer of 64) and $\geq$ 8.0 $\log_2$ (arithmetic titer of 256) to RSV A and B respectively are protective against RSV-associated-hospitalization; levels that were achieved by majority of older children and adults by natural RSV infection.

Another finding of interest is the observation that the serum antibody responses are closely linked to the genotype causing the infection [105, 106]. Due to the antigenic diversity it may be that if the immune response is group specific, infection by one group could occur despite the presence of neutralizing antibodies specific to another group [107, 108]. This then may explain general re-infection, re-infections in infants with maternal antibody and the yearly variation seen in the dominant genotypes isolated. The issue of cross immunity has not been comprehensively explored and forms the basis of the current RSV birth cohort.

2.3.1 Viral characteristics that permit re-infection

The seasonal shift in the dominant subgroup (or group) suggests that antigenic variation may play a role in the ability of RSV to escape the immune response and establish re-infections. Mufson *et al* [81, 109] assessed the group characteristics of the viruses which caused re-infections in 13 children. Of the 10 children with initial group A virus infections, 6 had group B viruses and 4 had group A viruses upon re-infection representing more group

B virus infections than would otherwise have been expected by chance, given the ratio of circulating groups during the study. The conclusion proposed is that an initial infection with a group A virus provided some degree of protection against re-infection by group A viruses. This conclusion is undermined by the fact that Group B was the dominant variant during the period of reinfections. Waris *et al* [44] observed that children older than 6 months during their first RSV infection were more resistant to homologous than heterologous group re-infections. A recent study by Scott *et al* [110], using samples collected within the main birth cohort project, identified 12 cases of re-infections. Out of the 12, 6 infants were infected in two consecutive epidemics. 4 were infected with RSV-A in the first epidemic followed by RSV-B in the second epidemic while the 2 others were infected with RSV-A during both epidemics. The RSV-A for these 2 infections had no significant G gene sequence variability between samples. A similar result was documented by Hall *et al* who showed that re-infections (that are very close in time) could occur by repeated exposure to the same viral isolate, so that antigenic variation was not strictly required to allow re-infections [101]. Two other infants in the Kilifi study were infected and subsequently re-infected with different RSV-A strains during the same epidemic. In the remaining 4 cases viral persistence was suspected, although re-infection with the same variant could not be ruled out. Data from this study suggests that after primary infection, some infants lose strain-specific immunity within a period of 7–9 months (i.e. between epidemics), while others lose group specific immunity and become infected with the same group (different variant) within 2–4 months (i.e. within an epidemic).

2.4 Infection and re-infection

To date, numerous studies have been carried out [4-9, 11, 27] and the literature is replete with evidence to indicate that RSV is a major cause of serious life threatening disease of the lower respiratory tract during early life.

i) Infection in children

In general approximately 50% of children contract their first RSV infection during their first year of life and by the time they are three years old the majority of children have experienced at least one RSV infection [12, 90, 111]. A summary of some reported infection rates in children is given in Table 2.3. Though relatively uncommon in the first month of life, a study carried out in Chile [112] reported that 28% percent of the cases observed were in children who were less than 1 month old and infection is frequently very severe if it occurs at this age [94, 113].

Table 2.3. RSV infection by age in young children from studies in developed and developing countries

Location (Reference)	*Age of children*	*Risk of infection (proportion with evidence of infection)*	
		Virus isolation	Serology
Kilifi [5]	0-11 months	39%#	
US [19]	2 month- 3 years	98%* 75%$ 65%&	
Houston, Texas [12]	0-11 months	68%	
	13-24 months	82.6α	
Sweden [114]	0-18 months	87%	
	0-3 years	100%	
Tecumseh, Michigan [115]	<1 years	14%	17%
	1-2 years	17%	
	3-4 years	26%	
	1-4 years		17%
Chile [112]	0-5 months	46%	
	6-11 months	35%	
	12-33 months	25%	

* first $second and &third infections
# 9% re-infections
α rate /100 cy

ii) Symptomatic versus asymptomatic infection

RSV infection is apparently rarely asymptomatic in primary infection or infection in the first year of life and results in disease in young children with a wide spectrum of symptoms [18, 55, 101, 116, 117]. Evidence from several studies show that RSV infection in older children and adults is associated with mostly mild illness [15, 16, 57, 58, 101, 118, 119]} though may sometimes be asymptomatic [101, 120]. In some cases it has been shown to be associated with a moderately severe or severe illness [101, 121, 122]. Several adult challenge studies have been carried out (Table 2.4). The adult multiple challenge study by Hall *et al* [101] reported a decline in symptomatic infections with each subsequent

challenge; 85% of the infected subjects developed symptomatic infection on first challenge compared to ≥50% of infections in the subsequent challenges being asymptomatic. Contradicting this are the results from the family study by Hall *et al* [15] in which they reported little difference in the type of illness or frequency of symptoms according to age. This may be related to the infecting dose known to be higher within the household setting [123] possibly resulting in more cases of symptomatic infection although it may also probably be a factor of the sample size. In this study there were no cases warranting hospitalization. Cases warranting hospitalization would be more likely to be in the primary/youngest individuals. However, studies of nosocomial infection (also presumably constituting higher infecting dose) have shown that 86% of infected adults will have a symptomatic infection and in some cases may be protracted [122, 124].

Findings from another adult challenge study corroborate this hypothesis of dose-dependent severity. Even though extensive infection occurred at both low and high dose levels (500 and 100,000 pfu), it was observed that illness only occurred following administration of the high dose of virus [99]. Illness was observed at levels of viral excretion less than or equal to those observed in the low dose study suggesting that the temporal pattern of viral replication is an important determinant of illness. It was suggested that illness seemed to be associated with a particular pattern of virus multiplication (rapid production of moderate amount of virus at high dose versus delayed replication observed at the low dose but ultimately resulting in greater total amount of virus). This rapid production therefore seems more effective at inducing illness.

Table 2.4. Adult challenge studies showing infection and illness rates after RSV infection

Study	*Population*	*Presence of antibody before Infection*	*Infections*	*Illness*
[58]	41 adult males volunteers	Neutralizing antibody pre-challenge	33	20
[99]	16 adult males low dose 500pfu	Neutralizing and secretory antibody pre-challenge	16	0
	17 adult males high dose 100000pfu	Neutralizing and secretory antibody pre-challenge	9	5
[101]	15 adults repeatedly challenged after natural infection	1)Natural infection	15	15
		2)Challenged at most 6 times	15	51% of total challenge infections

This issue of symptomatic versus asymptomatic infections is important insofar as it suggests that surveillance based on clinical signs only may not be effective at identifying most and/or all cases of infection; an issue that is potentially critical in this study. Viral-specific antibodies in blood or saliva can thus provide information on the infection status of the host in the absence of clinical illness, and are particularly of use in support of clinical surveillance information and will be used to compliment clinical surveillance.

2.5. Severity of RSV infection in children

The risk of developing lower respiratory tract disease (LRTD) following primary infection is between 25- 40% [12, 18, 19, 55, 96]. This risk has been shown to be age related in most hospitalization studies [37]. Evidence for declining risk of disease has also been shown in community studies. Data from the community study in US by Glezen *et al* [12] showed a decline in risk from infants upwards; 22.4/100 child-years to 7.7/100 child-years in children 37-48 months. The published data from Kilifi [5] showed no significant age-related

variation in risk in the first year of life, but more recent data (unpublished, DJ Nokes) suggests highest risk in infants (particularly children aged 3-5m) declining into the second and third year of life. This decline in disease rates in primary infection with increasing age is most probably a reflection of increased airway size of the infant but may also reflect increased immunocompetence. Disease burden estimates from different studies are covered in detail below.

Community versus Hospital infections: Incidence data from Developing Countries

Data from community studies are relatively scarce. The majority of the information is from hospitalised children. There is little that can be inferred about what is actually happening in the community as studies performed on admitted children, whose selection is based on severity, only represent a partial view of the impact of RSV in the community.

The BOSTID studies reviewed the role of viruses in causing ALRI in developing countries [27]; 70% of the ALRI were attributable to viruses and RSV was found to be the major cause in children less than 5 years of age. Few studies since then have attempted to quantify the incidence of RSV-LRTI in developing countries. These studies are not exactly comparable because of several reasons; different populations studied (hospital (Table 2.5) vs. community (Table 2.6)), differing study designs and surveillance methods used (column 3 and 4) as well as the definition of LRTI employed by each study (not included in table). From these studies the incidence of RSV-associated LRI is quite variable and ranges between 8/1000 cyo - 220/1000 cyo. Other factors that might contribute in part to this variation include the frequency of sampling, reason for sampling (i.e. whether or not clinical symptoms had to be present), the type of samples collected (Nasopharyngeal aspirate vs. nasal wash or swab) and the method of determining that an infection had

occurred. Laboratory diagnosis on the basis of immunofluorescent antibody test alone, the favoured method of RSV diagnosis in children [125], is still not as sensitive as that by assay combinations (serology and immunoflourescence) or molecular methods (PCR) [5], although is more sensitive than older methods such as the complement fixation text and some enzyme imunoassays [28]. Some of these influencing factors have been included in the Tables and are also important practical considerations in the intended study (Chapter 3). Studies adopting similar designs report relatively similar rates [5, 6, 11]. The reported rates from the studies in Kilifi and Columbia were arguably on the higher side possibly due to the short duration of studies in addition to the method of surveillance employed (combination of active and passive).

More recently as a result of a WHO initiative to generate more evidence on the role of RSV in causing LRTI, 4 studies arising out of (but not adhering to) the same generic protocol were undertaken [11, 117]. From these studies the incidence of RSV-LRI in children < 5 years was 34/1000 in Indonesia and 94/1000 in Nigeria while the incidence of severe RSV-LRI per 1000 child years was 5 in Mozambique, 10 in Indonesia and 9 in South Africa with majority of RSV cases occurring in infants. It seems decidedly necessary to generate such epidemiological data prior to starting any immunization programmes. Moreover, it would be important to compare incidence data from Kenya to that generated from other countries, both developed and developing, to identify similarities and any potential differences that may require further investigation.

Table 2.5. Incidence rates of RSV-LRTI in developing countries in hospital studies (adapted from review by Nokes [38])

Location, population (Reference)	*Period of study*	*Surveillance*	*Specimen details, diagnostic tests*	*Denominator Population (n) (Age group)*	*RSV incidence/ 1000cyo (95%CL)*
Kilifi District, Kenya Rural [5]	15 months 2 epidemics	Active home visit (weekly in season monthly otherwise), passive referral	NW clinic by TFW or CO IFAT	311 cy n=338 (<1 years)	13 (0-25)
Manhica, Mozambique Rural [11]	12 months 1 epidemic	Passive surveillance at IP of District hospital	NW in OP Rapid Ag ELISA	1342cy (<1 years)	15 (8-22)
		Passive surveillance at IP of District hospital	NW in ward Rapid Ag ELISA	6020cy (<5 years)	5 (3-7)
Lombok, Indonesia Rural [8]	24 months	Passive surveillance by village HC and hospitals. Effective referral procedures	NW at hospital Rapid Ag ELISA	15000cy n=15000 (<1 years)	25 (22-28)
				30000cy n=30000 (<2 years)	14 (13-15)
Western Region, The Gambia [9]	36 months 3 epidemics	Passive surveillance at 3 main hospital serving the community	NPA in ward IFAT	(<1 years)	18
Rural				n=20338 (<1 years)	9
Urban					
Soweto, South Africa Urban poor [41]	13 months 1 epidemic	Passive referral Hospital IP surveillance	NW in ward IFAT, shell vial culture	24000cy n=24000 (<2 years)	4 (3-4)*

CO- Clinical Officer HC- Health Centers
*severe RSV-LRTI

Table 2.6. Incidence rates of RSV-LRTI in developing countries in community based studies (adapted from review by Nokes [38])

Location, population (Reference)	*Period of study*	*Surveillance*	*Specimen details, diagnostic tests*	*Denominator Population, n Age class (yrs)*	*RSV incidence / 1000 cy (95%CL)*
Rio de Janeiro, Brazil Urban poor [4]	36 months 3 epidemics	Active weekly home visits	NPA (shallow) at home by HCW. IFAT, culture	786cy n=262 (<5 years)	14 (6-22)
Cali, Colombia Urban poor [6]	17 months	Passive referral, Active weekly home visits, regular clinic appointments, clinic referral for ARI	NPA in clinic IFAT and culture	413cy n=340 (<1.5 years)	200 (149-241)
				399cy n=205 (<1 years)	220 (156-284)
Ibadan, Nigeria Peri-urban and rural [11]	24 months 2 epidemics	Active weekly home visits	NW at home by CN Rapid Ag ELISA	1579cy (<5 years)	116 (78-154)
				316cy (<1 years)	94 (79-109)
Kilifi District, Kenya Rural [5]	15 months 2 epidemics	Active home visit (weekly in season monthly otherwise), passive referral	NW at home, clinic or ward by TFW or CO IFAT	311cy n=338 (<1 years)	154 (111-198)
Takhli District, Thailand Rural [126]	28 months 2 epidemics	Active surveillance for LRI with referral to hospital	NPA at hospital IFAT	14569 n=6244 (<5 years)	8 (7-10)

Bandung, Indonesia Peri-urban + rural [11]	*24 months 2 epidemics*	*Active weekly home visits, LRTI cases referred to clinic*	*NW at clinic by physician. Rapid Ag ELISA*	*284cy (<1 years)*	*41 (17-65)*
				1420cy (<5 years)	34(24-44)
Metro manila, Phillipines Urban poor [127]	24 months	Weekly home visits, LRTI cases referred to clinic. Passive referral to hospital	NPA at hospital. IFAT, culture.	1418cy (<5 years)	28 (19-37)
Kilifi District, Kenya Rural [5]	15 months 2 epidemics	Active home visit (weekly in season monthly otherwise), passive referral	NW at home, clinic by TFW or CO IFAT	311cy n=338 (<1 years)	100 (54-135)*
Bandung, Indonesia Peri-urban + rural [11]	24 months 2 epidemics	Active weekly home visits, LRTI cases referred to clinic	NW at clinic by physician. Rapid Ag ELISA	284cy (<1 years)	16 (1-31)*
				1420cy (<5 years)	10 (5-15)*

TFW –Trained field worker CO- Clinical Officer
* severe RSV-LRTI

It seems probable that socio-economic or cultural risk factors for infection/disease differing between communities may contribute in part to some of the differences observed, a subject that is explored in the current study. Although the findings from these studies are variable and to the greater extent underestimates (i.e. do not capture all cases of disease), these results make a substantial case for a significant burden of RSV-LRI. Some studies in the developed world [12, 13] have reported rates similar to those discussed above.

Incidence rates of RSV-LRI in the Developed World

The rates of severe illness among children < 1 year in developing countries [11] are comparable to the reported hospitalisation rates in developed countries (Table 2.7). The incidence of RSV-associated lower respiratory tract infection in developed countries varies from 3 to12 per 100 children in the first year of life [113]. Two community studies have reported incidence rates similar to those detailed in Table 2.6. A birth cohort study of low income families in the US reported an incidence rate of 224/1000 child years in infants [12] similar to reports from Kenya and Columbia. The second study of out patient presentations of LRTI in infants reported an incidence of 44/1000 cy [13] similar to reported rates from OP surveillance in Mozambique [11] and Columbia [7]. The majority of these RSV associated lower respiratory tract diseases occur in infants younger than 12 months of age with the most severe cases occurring in infants younger than 6 months of age with a peak incidence between the age of 2 and 6 months [90, 94, 113, 128].

Table 2.7. Incidence rates of hospitalization for RSV among children < 1year in studies from industrialized countries

Location, Reference	*Surveillance*	*Specimen details, diagnostic tests*	*RSV incidence / 1000 cy*
United States [12]	Longitudinal follow up of birth cohort	NW at ARI culture Serum Microneutralization assay	16
Austria [129]	Population-based retrospective review of admission over 1 year	NPA IFAT	6
England [130]	IP admissions-primary and nosocomial infection of city children	NPA, cough and nasal swabs IFAT and culture	20
Germany [131]	Hospitalized children with ARTI	NPA Multiplex RT-PCR	12 *
United States [13]	OP clinic	NW Culture	8
Norway [132]	Hospital admissions with ARI	NP sample Rapid IF and culture	10
Switzerland [133]	Hospital admissions	NW DIA, indirect ELISA confirmed by culture	5

NPA-Nasopharyngeal aspirate
RT-PCR Reverse Transcriptase Polymerase Chain Reaction
*Children <2 years

2.6 Re-infection

These data presented above highlight the substantial burden caused by RSV disease in young children resulting mainly from primary infection. RSV is also known to re-infect repeatedly throughout life. Appreciable attack rates in older age groups was suggested by Berglund [50] and later documented in the Seattle and Tecumseh studies [16, 134]. In the family study conducted after these by Hall *et al* [15], substantial attack rates were seen in all age groups; 16.8% in adults compared to 29.4% in infants. Results from this study

showed that age was not a primary factor in determining the attack rate (infection rates were substantial in all ages groups). Following these studies, two other studies involving young children also documented significant re-infection rates in children [12, 19]. Subsequently, Hall *et al* in repeat challenge studies in adults documented the frequent and potential ease of repetitive infection even with the same viral strain [101]; 73% of adults in this study were re-infected two times or more after a natural infection. More recently, several studies (reviewed below) have shown RSV to be an important cause of respiratory tract infection in adults and especially in the elderly [135, 136]. A summary of reported re-infection rates is given in Table 2.8.

The majority of these studies used serological testing to define infection rates and was carried out in developed countries. Mild illness in adults and older children indeed point to the potential benefit of a combination of methods; antigen positive and serology for the identification of infection in this population. Diagnosis of RSV during acute infection in adults is difficult because of the poor sensitivity of viral culture and antigen detection. Thus, a combination of diagnostic methods including molecular based techniques would enhance the diagnostic yield and provide a more accurate picture of the epidemiology of RSV in the present study.

Table 2.8. Estimates of risk of RSV associated re-infections

Study Area (Reference)	*Population*	*Study type*	*Age bracket*	*Infection History*	*Risk of re-infection (by serology or viral isolation)*
Washington DC [19]	Infants and young children	OP/ IP surveillance	infants and young children	previously known to be infected	75% (second infection)
				previously known to be infected	65% (third infection)
Chicago [18]	8 years, 11 patients	OP/ IP surveillance	6 children <1 year	previously known to be infected by viral isolation	100%
			5 children 1-3 years	previously known to be infected by viral isolation	100%
UK [20]	School children, 121	Community	7-10 years	pre-existing IgG Ab	18% (4-fold rise in IgG Ab)
					24% (2-fold rise in IgG Ab)
Nursery studies					
US [119]	Nursery school children, 16-24 observed for 41 weeks	Community	young children 3-5 years	15 children with neutralizing Ab	50% (virus isolation with symptoms)
US [137]	Nursery school children, 51	Community	young children	51 children with pre-infection Ab >= 1:4	66%
Family studies					
US [15]	36 families, 188 members in the US	Community	2-<5 years		26%

			5-<17 years		19%
			17-45 years		17%
Tecumseh, US [16]	1 year follow up of families	Community	1-4 years	Detectable CF Ab pre-infection	16% (significant rise in CF antibody)
			5-9 years		20%
			10-14 years		17%
			15-19 years		10%
			20-29 years		6%
			30-39 years		4%
			40-49 years		3%
			50+		-
Tecumseh [138]	6 years, Families recruited for 1 year follow up	Community	1-4years	Detectable CF Ab pre-infection	17% (significant rise in CF antibody)
			5-9 years		20%
			10-14 years		17%
			15-19 year		10%
			20-29 years		8%
			30-39 years		4%
			40-49 years		7%
			50+		7%
Houston, Texas US [12]	Families, 92 children	Community	13-24 months	previously known to be infected by viral isolation	76%
			25-36 months		45%
			37-48 months		33%
			49-60 months		50%

Finland [50]	1 season family members of child with RSV infection	Community	siblings (brother & sisters)	significant rise in CF Ab	23%
			adults family members		16%

Ab- Antibody
CF Ab - Compliment fixing antibody
OP- Out patient
IP- Inpatient

These findings suggest that re-infection is common although disease is reduced (discussed below) indicating that natural immunity is incomplete. Factors that may contribute to the occurrence of re-infections include the time interval between infections or viral challenge [101], antigenic variation of the virus (covered in detail in section 2.3.1), age and serum neutralizing antibody response of the host (section 2.3) as well as the duration of acquired immunity. The frequency of RSV re-infection throughout life seems to indicate that a large susceptible proportion is consistently available. It seems probable, given the epidemic pattern observed, that once infected individuals have some immunity resulting in the depletion of these susceptible individuals. This acquired immunity is then lost relatively quickly allowing the virus to invade the population resulting in another epidemic. These re-infections in older individuals may be the primary source of serious infections in infants and young children. Indeed some studies have shown a link between school children and introduction of infection into the home [15, 55, 117, 128, 139]. This contribution of re-infections in older children and adults in the developing world and to the observed pattern has not been evaluated. The household study presented here was therefore carried out with the aim of describing community re-infection patterns and illuminating the possibility of viral persistence in the off-season using regular oral-fluid sampling (see chapter 4).

Infection in Adults

In contrast to studies in children, few studies have looked at RSV infections in adults. RSV infection in adults, though previously recognized, has largely been overlooked because of the significance of RSV in children. In adults the rate of infections (re-infection) is about 5% per epidemic [16, 134]. Studies carried out in the United Kingdom and in the United States, have reported incidences of increased antibody titre to RSV of between 17 and 19 % in older individuals during an epidemic [16, 20]. In other studies involving surveys of

serum from adults, it has been shown that 33-99% of individuals possessed compliment-fixing antibody (marker of present or recent past infection) for RSV [140, 141]. Recent findings from a study by Hall *et al* [65] in the adult population suggests that RSV re-infection in healthy working adults is an under-recognized cause of morbidity. Healthy subjects 18 to 60 years of age were evaluated for respiratory virus infection over a 20 year period through a surveillance program. 7% of total number of subjects studied experienced an acute RSV infection. The infections were symptomatic in 84% of subjects, involved only the upper respiratory tract in 74%, and included lower respiratory tract symptoms in 26%. A similar community-based observational study from the UK whose focus was also on the burden of RSV disease in healthy, working adults [142] also found a significant burden of RSV infection in adults. Individuals presenting with an influenza-like illness between 1995 -1998 were evaluated. RSV was isolated from 20% of patients between 15 to 44 years of age. These findings show that the health care burden imposed by RSV on healthy working adults is not negligible and that quite substantial proportions are symptomatic. As noted, all these studies have been in the developed countries and as yet there is no data from developing countries on the role of RSV in older members of populations. This study will aim to quantify the community rates of re-infection in older children and adults. It would be expected that nearly similar rates of infection as discussed above would be observed in the Kilifi population; also classified as generally healthy.

2.7 Household Studies

Family studies also provide an ideal opportunity to study RSV re-infections in older age groups. Family studies conducted by Fox *et al* [143] showed that illnesses with a respiratory component comprised 78% of the total sickness observed. RSV infection has since been shown to spread within families with a substantial proportion of children as well

as adult family members contracting infection [15]. In this prospective study, siblings were found to be the most common primary case in the family setting [15]. Approximately 40% of all family members over one year of age were infected. This family occurrence seems to be one of the epidemiological characteristics of RSV infection [15, 50] since it requires close contact for spread implying that social situations existing in home and school settings are important for transmission. Indeed respiratory infections are often carried home by a school-age child and passed onto a younger child, especially an infant [15, 16, 128, 144]. Despite the potential importance for transmission and disease, little is known about the occurrence of RSV infections in families and about intra-familial spread of those infections (Table 2.9). No such studies have been carried out in the recent past.

Table 2.9. Summary of studies of defining respiratory syncytial virus infection within families

Studies	*Study Design*	*Age Group*	*Proportion Infected (by virus isolation [a] or serology [b])*	*Other Observations*
Rochester, US 36 families, 188 members in the US [15]	Families with 2+ children, one < 1yr. Visited every 3 to 4 days for two months. Nose & throat specimens collected from all, symptoms documented.	< 1	29%[a]	Average number of children / family 3.5 School child most likely introducer
		1-<2 years	29%	
		2-<5 years	26%	
		5-<17 years	19%	
		17-45 years	17%	
Tecumseh, US 1 year follow up of 341 families [16]	Families under surveillance of acute illnesses. Blood samples were obtained from family members at 0, 6 & 12 months.	1-4 years	16% [b]	High frequency of infection in school - age children
		5-9 years	20%	
		10-14 years	17%	
		15-19 years	10%	
		20-29 years	6%	
		30-39 years	4%	
		40-49 years	3%	
Tecumseh , US 6 years of data [138]	Families recruited for 1 year follow up each.	1-4years	17% [b]	
		5-9 years	20%	
		10-14 years	17%	

		15-19 year	10%	
		20-29 years	8%	
		30-39 years	4%	
		40-49 years	7%	
		50+	7%	
Houston, Texas US Families, 92 children [12][c]	Birth cohort. Families visited weekly in season, all members examined at time of acute illness. NW obtained at time of each contact. Serial serum samples obtained from index patient - 4 samples in 1st yr & 2 spanning each subsequent RD season.	13-24 months	76%[ab]	
		25-36 months	45%	
		37-48 months	33%	
		49-60 months	50%	
Finland 1 season, family members of child with RSV infection [50]		siblings (brother & sisters)	23%[ab]	Mothers had higher infection rates than fathers. Children most likely introducers
		adults family members	16%	

In the Tecumseh sero-epidemiological study, a serological response to RSV was seen most often in school children aged 5-9 years old [16]. 20% of them were infected during each yearly epidemic. The rate of re-infection was observed to fall rapidly with increasing age during childhood and then more slowly among adults (Table 2.9). Contrary to the observation by Hall *et al* [15], in this study differences in the spread of RSV within different families were noticed. As the number of members in the family increased from three to six people, there was an increase in the proportion of families in which RSV infected one or more individuals; 5.5% in families with 3 members vs. 16.3% in families with six members. This is most likely due to the additional opportunity for introduction of the virus with increasing family size. Larger families also experienced more multiple infections once the virus was introduced. As was the case in the Rochester study [15], data from this study supported the idea that it is the school-aged child who more often than not introduces the infection into the family unit. Infection and illness rates were also found to be higher in mothers than fathers as presumably the mothers had more intimate contact with the children than did the fathers [50]. A similar observation (increased number of ARI episodes in mothers) was seen in the Cleveland family studies [145]. This observation seems to suggest that virus dosage and exposure time may play a part in the development of RSV infection in adults. All these studies have been carried out in developed countries, countries whose social structures are different to those in developing countries. Thus the current study provides an ideal opportunity to examine infection within families and generate information on infection in the community in rural Africa.

2.8 Risk factors for ARI and ALRI

Comparative studies have shown that while the incidence of ARI in developed and developing countries does not differ markedly [16, 115, 146-152], the case fatality rates in

children < 5 years is 5-10 times higher in developing as compared to industrialized countries [153]. The excess mortality rates from ARI in the developing world suggest that several risk factors may be responsible. If these risk factors in developing countries predisposing children to severe disease after initial infection can be identified and prevented, much of this mortality burden due to ARI can be prevented in developing countries as well.

Factors that increase the risk of acquiring disease

It is known that certain risk factors increase the likelihood of a young child or infant acquiring a respiratory infection consequently increasing the risk of severe respiratory disease. These risk factors for ALRI can be categorized into (i) those which act at the individual level which are child specific and include birth factors (birth weight, gestational age), breastfeeding status, nutritional status and some immune factors (antibody levels, infection history) and (ii) those at the household level which include parental characteristics (education, age), household size and income, crowding and indoor air pollution [154]. A summary of risk factors for ARI and ALRI as described in the literature is presented (Table 2.10).

Table 2.10. Factors associated with ARI and ALRTI: summary of findings form various studies

Factor	*Findings and comments*	*Reference*
Household characteristics		
Household income level	Association between a low level family income and ARI in developed countries.	[115, 146-148]
	Similar relationship from studies in developing countries	[155, 156]
	Higher income level is usually associated with better general health, less crowding index	
Parental characteristics		
a) Maternal Education	Conflicting results - no clear established relationship.	
	Low parental literacy was a significant risk factor for ALRI	[157]
	Mother's education was also inversely and linearly associated with ALRI mortality	[155, 158]
	Maternal education was not associated with ALRI mortality or with the incidence of ARI or ALRI	[27, 156, 159]
b) Maternal age	Differing results on maternal age as a RF	
	Association between maternal age and ALRI mortality	[155, 156]
	No consistent association observed between maternal age and incidence of ARI or ALRI	[27]
Crowding and Family Size	Crowding generally enhances the spread of respiratory agents	
	Increased incidence of ARI with crowding in studies done in both developed and developing countries	[156, 157, 160-162]
	The presence of acute respiratory tract disease in other household members was a significant risk factor for acute lower respiratory tract infection in children OR 5.6 (IP) and 5.2 (OP).	[163]
Smoking	Various studies have consistently shown an increase in the risk of ARI with smoking	[155, 156, 164, 165]
Indoor Air	In developing countries, pollution of the indoor	

pollution	air from biomass fuel is quite high	
	Conflicting results	[166]
	Several studies from developing countries have shown associations of ALRI with indoor air pollution in the crude analysis other studies have found no association	[163, 166-169]
	Few studies have adjusted for confounders-	[170-172]
	Increased risk in children in developing countries who were carried on their mother's back during cooking	[164, 173].
<u>Child Level factors</u>		
Malnutrition	PEM has consistently been reported as a determinant of ALRI and ALRI mortality from studies conducted in several countries	
	Strong association between moderate/severe protein energy malnutrition (PEM) and incidence of Pneumonia	[7, 174, 175]
	Linear association between weight-for-age z-scores and/ or weight-for-height and height-for-age and ALRI and/or ALRI mortality	[27, 155-157, 159, 168, 176, 177]
Breastfeeding	Studies have generally shown a protective role of breast feeding against infection	[178]
	Significant association, increased rates of pneumonia/ALRI and lack of breastfeeding	[175, 179]
	Breast feeding was shown to provide some protection for pneumonia and/or ALRI mortality	[180, 181]
	No protective effect of breast feeding against ALRI mortality	[156]
	Breastfeeding has for many years been recommended as being good for child health.	
Established determinants	Increasing age, female sex, birth weight over 2,500 grams, and immunization against childhood disease (like measles, Hib) are factors that are known to be protective against ARI and ALRI	[160, 161, 175, 179]

PEM –Protein energy malnutrition; RF- Risk factor

The observed difference in the effect of various risk factors may be a reflection of the differing study designs (case-control or cohort) and populations studied (hospitalized cases vs. milder community infection) and differences in the methods of analysis used- some studies adjusting for potential confounding [163, 164, 166, 172, 173] whilst others reporting crude estimates. Moreover, some of the studies reported had a retrospective data collection design and were thus susceptible to the usual methodological limitations of such studies (for example recall bias). Recall accuracy declines as the level of detail requested increases. Nonetheless, there is evidence from various studies supporting the effect of different factors; socio-economic classification, crowding, malnutrition and exposure to smoke. PEM has consistently been reported as a determinant of ALRI and ALRI mortality from studies conducted in several countries over the past two decades. This underlines the importance of considering this variable as a potential confounder in any evaluation of ALRI and potential risk factors. There are several interventions some as simple as breastfeeding that potentially alleviate malnutrition as a risk factor. Data on established determinants such as increasing age, female sex, birth weight over 2,500 grams, and immunization are individual factors that are known to be protective of ARI and ALRI [160, 161, 175, 179] and need to be included in ALRI risk factor studies to adjust for them in the analysis.

There is limited data on the prevalence of concurrent infections in cases of LRTI and their potential effects on the disease manifestation. A study of co-morbidity in childhood in Africa [182], found significant co-occurrence of infection in the community and evidence of co-occurrence of diarrhoeal diseases and pneumonia, with greater co-morbidity with increasing severity of disease. However, there was no evidence of a synergistic effect on mortality risk. There is potential in such cases for overlap between risk factors for the different infections or it may be that one infection predisposes or results in increased risk of

the other infection e.g. with Malaria which is thought to suppress host resistance to viral/ bacterial pathogens [182]. Still it is possible that this observed co-morbidity may be artefactual; in malaria endemic areas, the potential for an over-estimation of LRTI cases exists since the symptoms of LRTI can be consistent with malaria diagnosis [11, 183]. In the Kilifi study [5], 5% of children with LRTI also had malaria parasitaemia. No difference was observed in the occurrence of sLRTI or ARI mortality in an intervention trial in Gambia in two groups of young children one of which received malaria treatment [184]. In addition, viral pneumonia aetiology studies report the presence of more than one virus in a fraction of the cases with one study showing increased disease severity in cases of dual viral infection. Studies are indicated in developing countries to further investigate the prevalence of dual infections.

Risk factor for RSV Infection and Disease

Although risk factors for RSV infection and disease in developing countries are not well defined, differences in disease estimates (section 2.5) compiled from various localities [5, 6, 11, 176] point to the possibility of some underlying differences in risk factors. This creates a need to understand the underlying risk factor dynamics in these regions and thereby possibly explain the differences in observed estimates. Table 2.12 provides data from a number of studies mostly from developed countries.

Table 2.11. Factors associated with RSV infection and disease: summary of findings form various studies

Factor	*Findings and comments*	*Reference*
Gender	Boys suffer more severe disease than girls - Boys have shorter and narrower airways and therefore are more likely to have development of bronchial obstruction upon RSV infection	[94, 116, 131, 133, 185-188]
	An analysis of representative studies by Simoes found that the risk ratio of boys to girls was 1.425:1 (95% CI: 1.40, 1.45; range, 1.2 to 1.7:1).	[189]
Gestation/ Birth weight	Independent association of gestational age and birth weight and risk of RSV hospitalisation – more disease occurring in premature (i.e. low birth weight) babies.	[186]
Birth month and Age	Several studies show association between birth month and risk of RSV-LRTI and RSV hospitalisation Probably to do with presence of Maternal Ab i.e. births occurring 2-4 months before an epidemic are likely to be infected when MatAb are decayed but they are still very young.	[94, 113, 186, 190]
	Several studies have reported a relationship between age and infection with most disease occurring in infancy	[94, 132, 133, 185, 186, 191-193]
	i) Most disease occurs in the first few months of life so birth during the first half of the RSV season would be an obvious risk factor	[94, 189]
	ii) Also a reflection of the maternal antibody levels i.e. children born in the first half are born to mothers without recent infection history thus have lower antibody levels	
Breastfeeding	Contradictory results reported as far as breast feeding is concerned with some studies showing protection and some not	[188, 191, 194, 195] [189, 191]
	The biological processes proposed for protection included: i) Colostrum, which contains both anti-RSV IgA87 and lactoferrin, may have an important antiviral property.	[189]

	ii) ii) Breast milk promotes lung maturation, perhaps through prolactin	
Crowding and siblings	Most of the evidence seems to show that crowding is a risk factor.	
	RF for hospital admission - cases came from larger more crowded compounds and the risk was greater depending on the number of kids between 3-5 years in the homestead	[194]
	Significant effect of the number of persons sharing a bedroom	[113]
	Living with siblings in school was a significant risk for RSV hospitalization in infants.	[190, 196]
	Having 4 or more children under the age of 12 in the household was a significant risk factor for RSV hospitalization.	[191]
	Older siblings identified to be the most likely to introduce the virus into the family.	[15]
	Significant association between respiratory infections in infants and having at least one sibling especially if they were school age siblings.	[144]
	Presence of siblings in the family was found to double the risk of hospitalization	[197]
	Presence of siblings up to 5yrs older than the case was a significant independent risk factor. No effect with the presence of more than one older sibling or of siblings more than 5 years older than the case and also with crowding (m^2/resident)	[186].
	Lower rates of infection in infants without siblings and significantly fewer kids in households of controls.	[144] [198]
	Crowding is probably related to higher viral innoculum through interpersonal contact as well as to the increased probability of contact with an infectious person in a crowded environment	
Maternal Antibody	Observed a positive correlation between protection against infection and level of	[94]

	antibody	
	Mean titer of IgG higher in mothers whose children were uninfected	[93]
	Increased risk RSV-LRI in first year of life associated with low cord serum antibody.	[113]
	No association between the level of maternal neutralizing antibody and the risk of hospitalization	[191]
	Study of patients <1 yr (i.e. first infection) found no protective effect of either detectable IgG or neutralizing antibody to two major groups of RSV	[198]
	Case control study showed relationship between high Mat Ab and reduced risk of RSV disease	[199]
	Results from these studies are inconclusive.	
Malnutrition	Interestingly, several studies have found no association or negative association between malnutrition and RSV LRTI.	[45, 116, 188, 200-202]
	It has been suggested that malnourished children do not mount an exuberant immune response to RSV thus inadvertently protecting them from severe disease.	[189]
Smoking and Environmental Smoke	Conflicting reports on the effect of smoke exposure	
	Smoking during pregnancy was found to be an independent risk factor for hospitalization	[186]
	Significant relationship between RSV and smoking mothers at home	[144]
	No appreciable differences observed in smoking patterns in households in which cases of RSV occurred compared to those without cases of RSV infection.	[15]
	Controls were more likely to have been exposed to cooking smoke.	[194]
	Mode of action of smoke exposure unclear; is it increased risk of infection or is it an increase in	

	severity once infected.	
Co-infection	Evidence in the RSV literature for effect of co-infections is scarse.	
	Malaria	
	Mechanisms for positive and negative associations between malaria and RSV exist.	
	Malaria infection was supressed in children infected with other viruses	[203]
	2.3% of RSV cases in Kilifi (malaria endemic area) had concurrent malaria	[5]
	Placental malaria was associated with reduced transplacental transfer of RSV specific antibodies, potentially predisposing infants to earlier and more severe RSV.	[204]
	In Mozambique, a study of paediatric IP found a significant negative association between RSV associated LRTI and falciparum malaria-possible false association as the prevalence of parasitaemia in RSV cases was similar to that observed in community	[205]
	HIV	
	There is little data available on the association between HIV status and RSV infection/disease.	
	In paediatric admission with RSV positive sLRTI, mortality was higher (7.6 vs. 1.7%) and the prevalence of bacteraemia increased (15% vs. 3%), in HIV infected (HIV+) relative to HIV uninfected (HIV-) individuals	[206]
	The burden of RSV in the HIV+ population was observed to be significantly greater than in the HIV- population (14 vs. 3 cases sLRTI/1000) although the prevalence of RSV was lower in HIV+ compared to HIV- cases (in contrast to the relationship with bacterial pathogens.	[41]
	No association between HIV and RSV disease severity was identified in admissions <2 years old with LRTI. The population prevalence of HIV was low and majority (87%) of cases were non-severe LRTI which may explain the observed difference with previous study above.	[201]
	Other viruses	
	It appears that the epidemic behaviour of RSV	[9]

	can also be modified by interaction with other viruses	
	Dual infection with hMPV and hRSV confers a 10-fold increase in relative risk (RR) of admission to a pediatric intensive-care unit for mechanical ventilation	[207]
High risk groups	Studies in developed countries have identified several high risk groups in relation to RSV infection and hospitalization.	
	Increased risk of severe RSV infection and increased hospitalization in premature children, in children with CLD and CHD and history of BPD	[185, 191, 192, 197]
	Mothers of cases were more likely to be asthmatic.	[194]

CLD - Congenital Lung Disease, CHD- Congenital Heart Disease, BPD- Brocho-pulmonary disorder

Generally from these studies the main factors associated with increased risk of RSV disease or RSV hospitalization seem to be: primary infection at a young age, infection in children with underlying predisposing risk factors such as prematurity (presumably because of absence of maternal antibody), and cardio-pulmonary disease, male gender. Some environmental and social factors such as crowding, exposure to smoke or socio-economic classification have also been reported to increase the risk of acquiring disease.

Interestingly, unlike what is reported for ARI/ ALRTI several studies have reported no association or negative association between malnutrition and incidence of RSV disease [200, 208]. A negative association between RSV infection and malnutrition in children admitted with LRTI was observed in Mozambique [205], Indonesia [8], Nigeria [202], and Mexico [209]. A comprehensive review of the strength of evidence (mainly from developed countries) on risk factors for RSV-LRTI was done by Simoes [189]. The risk factors identified include: the male sex, age <6 months, birth during the first half of the RSV season, crowding/siblings and daycare exposure. The evidence for tobacco exposure was mixed while low maternal education (proxy for SES), lack of breast-feeding, and

malnutrition did not appear to increase the risk of severe RSV LRI. Few risk factor studies for RSV have been carried out in developing countries [126, 194, 200, 202] providing an ideal opportunity with this study to add to the body of knowledge. Furthermore, from the literature, risk factors for infection and disease are not differentiated thus there exists some confusion between the two. However, it is possible that the risk factors for infection and disease may well be quite distinct and this has not been assessed. The birth cohort provides a unique frame work for the comparison of risk factors for mild infection and disease as well as comparing risk factors for non-specific LRTI with those of RSV-LRTI. It has been shown that differences may exist (e.g. malnutrition) in the risk factors associated with the two outcomes. An attempt was therefore made in the present study to determine risk factors for clinical infection, RSV disease and compare these with those identified for non-specific LRTI (Chapter 8 and 9).

2.9 Potential use of Oral Fluid to determine infection rates

Antibodies in blood are useful as specific markers of immunity and past infection for various infections. Serological studies therefore have the potential to identify infections whether asymptomatic or symptomatic. However, it is not always feasible or desirable to collect blood samples by which to determine specific antibody status or titre. Less invasive biological samples for the detection of antibodies would and indeed have proved invaluable [210-218] in exploration of current infection (IgM, rising titre of IgG or seroconversion) or past infection and immunity (IgG). Anti-viral immunoglobulin M (IgM) is usually used as an indicator of recent infection [219, 220]. The presence of virus-specific immunoglobulin G (IgG) antibodies provides evidence of passive maternally-transferred antibodies or a past or present acquired infection in some infections is indicative of immunity (e.g. measles, rubella, but not in RSV).

Over the last few years various techniques have been developed to identify anti-viral antibodies in other body fluids [215-217]. Oral fluid samples offer a minimally invasive alternative to serum collection for the detection of virus specific antibody to common infections. It can thus be used as a means of determining incidence estimates. However, in infant (especially in first 6 months) serological response following infection is not always detectable [88, 90, 221-223] hence it not advisable for use as the only method of incidence estimation and thus would be supplementary or in addition to clinical surveillance. The collection of an oral fluid sample is easy, safe and is acceptable to the general public [213, 214].

Oral fluid is a mixture of several components. Most of antibody of diagnostic importance is present as a transudate from the capillary bed beneath the margin that separates the teeth and the gum known as the gingival crevice [224]. The antibody content of this crevicular fluid closely reflects the immunoglobulin class and specificities of the antibody found in plasma although saliva contains much lower concentrations of antibody than is found in blood [225] (Table 2.12). The concentration varies with the antibody component of the oral fluid sample tested which is determined by the way the sample is collected. The antibodies found in the crevicular fluid represent the composition of antibody in plasma including IgM, IgG and IgA class antibodies. Although the antibody concentrations fluctuate with the changing proportion of saliva and crevicular fluid [224], oral fluid is a good substitute for serum and can be used to determine recent, past or current infection.

Table 2.12. Mean immunoglobulin concentrations (mg/L) in different salivary components (adapted from Mortimer *et al [225]*)

Antibody Component	*IgG*	*IgM*
Plasma	14,730	1,280
Parotid Saliva	0.36	0.43
Crevicular Fluid	3,500	250
Whole Saliva	14.4	2.1

Because of the low and fluctuating levels of immunoglobulins in oral fluid, there is need for tests of high sensitivity, and which are standardized in some way against total specific-isotype level e.g. total IgG. Oral fluid assays are thus still less sensitive than serum assays. In spite of this, the advantages for the use of oral fluid clearly outweigh the disadvantages in circumstances where blood collection is undesirable e.g. outside the clinic setting. The process is minimally invasive, does not necessarily require professionally trained personnel, easy to use in children, and avoids needle stick injury and the inadvertent transmission of blood borne pathogens [212, 213].

2.10 RSV infection in Kenya

In Kenya, little information is available regarding the burden of RSV infection and disease. One study in Kenya in the 1980s, in the National Referral Hospital, revealed RSV to be the major viral pathogen accounting for approximately 20% of the hospitalised ARI cases [226]. To the researcher's knowledge no detailed surveillance of aetiology of ARI has been undertaken elsewhere in Kenya and no data is available on community level of transmission. In Kilifi district, prior to 2002, almost nothing was known about the importance of the RSV on population health, and at Kilifi District Hospital (KDH) little was known of the frequency of admissions of ARI due to RSV. From 2002 continuous surveillance was undertaken of RSV disease in child in-patients admitted with ARI to

KDH. The aim was to determine the burden of RSV disease in this patient population, the relationship of this disease to specific genotypes (variants of the virus), and the seasonal and spatial distribution of the cases in the community.

Children admitted to KDH with a diagnosis of ARI can be grouped according to severity (based on WHO guidelines [227]), into non-pneumonia, and mild, severe and very severe pneumonia. As the vast majority of RSV disease occurs in children under the age of 5 years [12, 19, 186, 228], sampling was restricted to children under 5 years, who were either admitted with severe or very severe pneumonia, or who on admission required treatment for LRTI. An additional inclusion criterion was children with a requirement for oxygen due to oximetry saturation reading <90% (hypoxia) as adopted in similar studies in The Gambia [116]. The graph below (Fig. 2.5) shows the outcome of surveillance at the KDH paediatric ward over four years between 2002 and 2005.

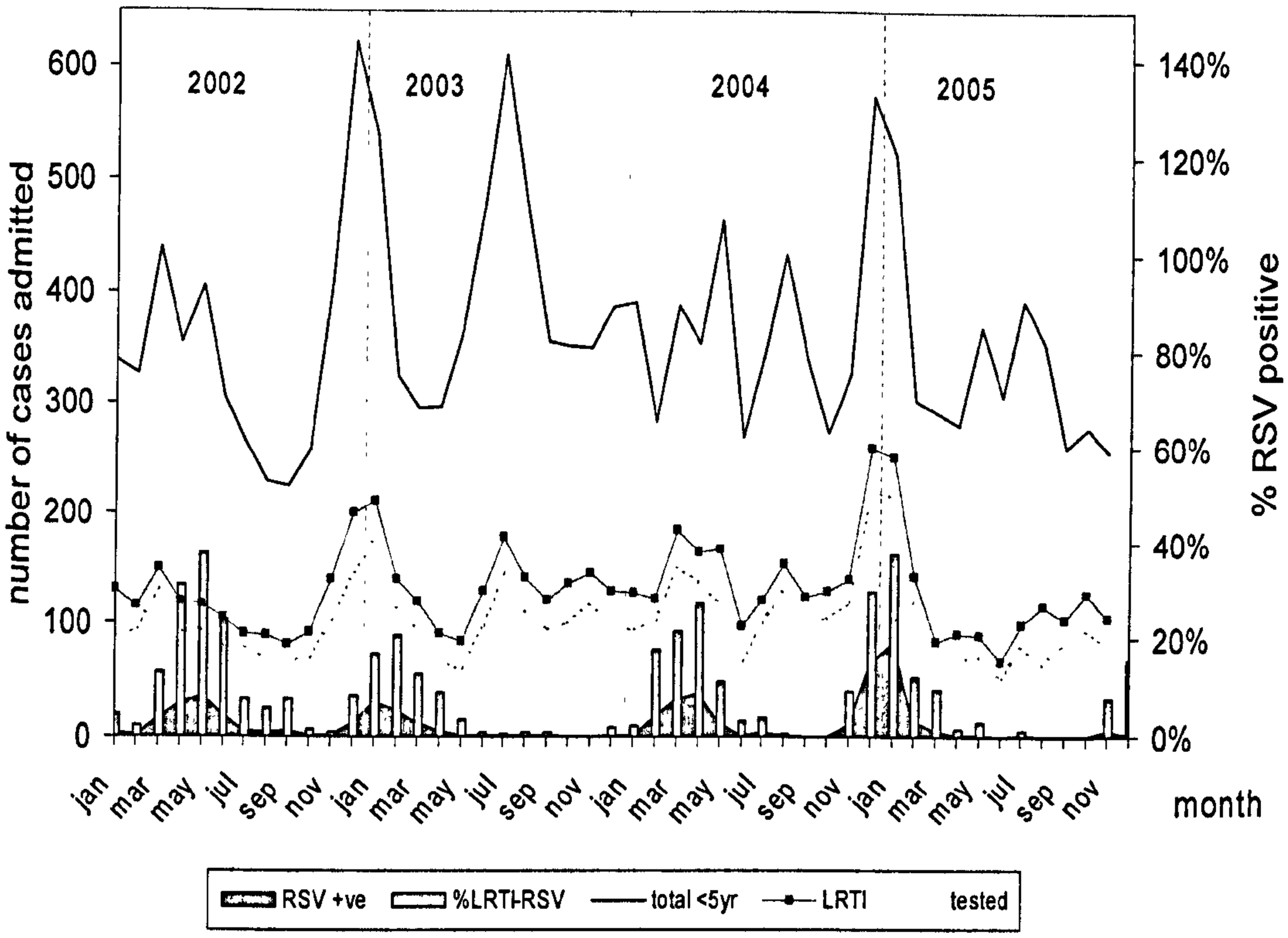


Figure 2.5. Kilifi District Hospital inpatient surveillance of RSV.

On average, approximately 400 children <5 years old are admitted each month (Fig.2.5 black line), with about one third of these having an admission ARI diagnosis (square markers). Of these, approximately 80% were sampled (nasal washing or nasal pharyngeal aspiration) and tested for RSV using an immunofluorescent test (IFAT) (dotted line). Of samples tested, between 10-40 samples were RSV antigen positive during the epidemics. While RSV accounts for only 3% of all admissions, annually, RSV accounts for some 10% of all ARI. Furthermore this burden is highly seasonal, and the proportion of ARI due to RSV can be as high as 40% during epidemics.

Chapter Three

Study Description

3.1 Overview

This chapter describes in detail the methodological approach adopted to collect field data in this study. The household study (results are presented in Chapters 5-7) was based on a cohort design and was nested within the RSV birth cohort. This entailed the intensive surveillance of a birth cohort of which there were two phases of recruitment staggered by one year (see Fig. 3.5 later). Household members of selected birth cohort children (those recruited in the second phase) were recruited into the household study and followed up over a period of two and a half years. The risk factor survey for LRTI and RSV-LRTI described in Chapter 8 was carried out within the main RSV birth cohort. These methods are described in the paper by Nokes *et al* [5].

3.1.1 Chapter Aims

The main objective of the study was to describe RSV transmission dynamics within the household and in the community. Specifically, this chapter seeks to give a general description of the study area and of the population studied. Details of the study design specifically describing the relationship of the household study to the birth cohort are given. The mechanics of field work, the enrolment, ethical issues, sampling, implementation, surveillance procedures, quality control (reliability) and data handling are detailed. Laboratory methods employed in the study are summarized and discussed and comparisons made between the methods used in this study and other options. Finally, some general results (i.e. % response, dropouts, surveillance intensity) are given as well as a description of epidemics in the main cohort and in the household cohort.

3.2 Study area

3.2.1 Geography

The study area falls within Kilifi district, a rural location in Coast province of Kenya.

Kenya spans the equator on the east coast of Africa between latitudes 5° north and 5° south. Kilifi district borders the Indian Ocean to the east, Mombasa to the south, Taita Taveta to the west and Malindi to the north (Figure 3.1.). Kilifi district covers an area of 4,779 km^2 with a density of 114 persons/ km^2. Kilifi town is home to some 36,000 residents out of a total district population of 544,303 people (1999 census) [229]. According to the Central Bureau of Statistics population projections, Kilifi district has a population of 679,499 in 2007. The equivalent for Kilifi Township is about 45,000 people. The district population consists predominantly of farmers, 80% of whom depend primarily on agriculture [230].

According to the Central Bureau of Statistics report (2003), Kilifi District is one of the least developed districts in the country with 72% of its people living in absolute poverty as measured by the head count index (i.e. the proportion of the population whose economic welfare is less than the poverty line, Ksh 1239 and 2648 for rural and urban households respectively, this is the expenditure required to buy a food basket that allows minimum nutritional requirements set at 2250 calories per adult person). The head count index for Coast province within which Kilifi district falls is 61% [231]. The literacy level in this district is 63% (males 76.9% and females 35%), is lower than the national average of 78% [232].

The population growth rate of Kilifi Town is put at 3.05% with a birth rate of 49.2 per 1000. The infant mortality rate for Kilifi District as reported by the Government is 85/1000 with an <5 mortality rate of 141/1000 [230]. The immunization coverage in the

demographic surveillance system (DSS-discussed below) stands at between 85% and 96% by both child health card records and history [233].

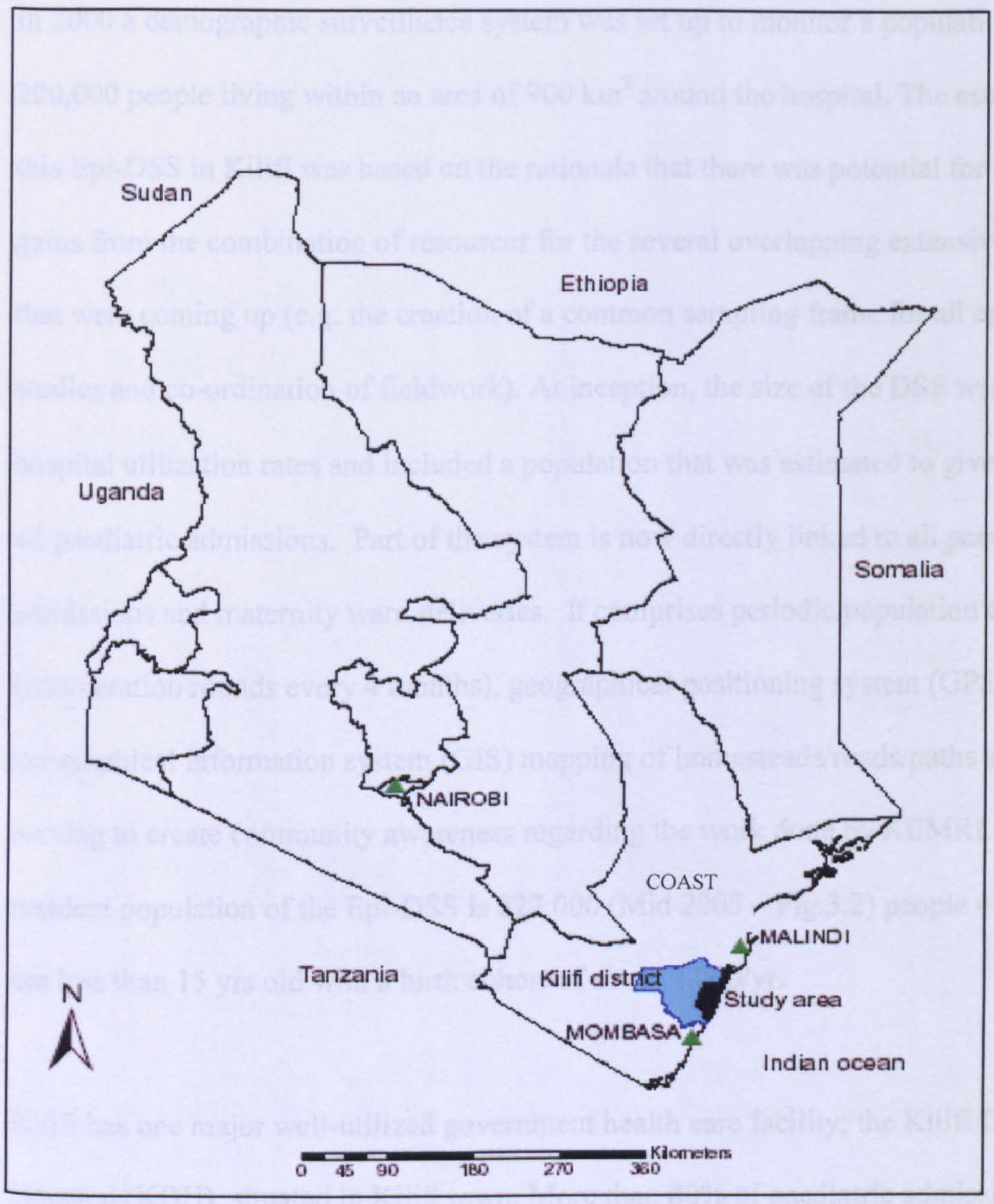


Figure 3.1. Map of Kenya showing the location of the study area (black) within Kilifi district (blue) set in Coast province (one of eight provinces in Kenya delimited by black lines).

3.2.2 Epidemiological-Demographic surveillance system (Epi-DSS)

The study was carried out in an area that was under continuous demographic surveillance. In 2000 a demographic surveillance system was set up to monitor a population of about 200,000 people living within an area of 900 km^2 around the hospital. The establishment of this Epi-DSS in Kilifi was based on the rationale that there was potential for synergistic gains from the combination of resources for the several overlapping extensive field studies that were coming up (e.g. the creation of a common sampling frame for all epidemiological studies and co-ordination of fieldwork). At inception, the size of the DSS was linked to hospital utilization rates and included a population that was estimated to give rise to 80% of all paediatric admissions. Part of the system is now directly linked to all paediatric admissions and maternity ward deliveries. It comprises periodic population census (enumeration rounds every 4 months), geographical positioning system (GPS)/ geographical information system (GIS) mapping of homesteads/roads/paths as well as serving to create community awareness regarding the work done by KEMRI. At present the resident population of the Epi-DSS is 227,000 (Mid 2005 – Fig.3.2) people 47% of whom are less than 15 yrs old with a birth cohort of about 7,000/yr.

Kilifi has one major well-utilized government health care facility; the Kilifi District Hospital (KDH), situated in Kilifi town. More than 80% of paediatric admissions to the KDH were found to originate from this DSS indicating good access to and utilization of the hospital. Approximately 15% of births in the DSS area occurred in KDH. KEMRI CGMR-C runs the KDH paediatric wards comprising at the time of this study a general ward of 40 beds with an occupancy usually > 100%, sometimes up to 200% during the malaria season and a 6-bed high dependency ward (occupancy between 50 -150%) where severely ill children are managed. It also manages an Out patient (OP) clinic within the hospital

grounds. The paediatric OP and IP services are supported by well equipped laboratories to undertake haematological, microbiological, biochemical and blood gas analyses. Malaria is endemic in the area, with peaks during the long (April–July) and short (November-December) rainy seasons [234]. A break down of hospital presentations from the resident population of the DSS is detailed in Figure 3.2. Overall, Malaria is the main diagnosis followed by LRTI. In the < 1 population LRTI is the main presentation.

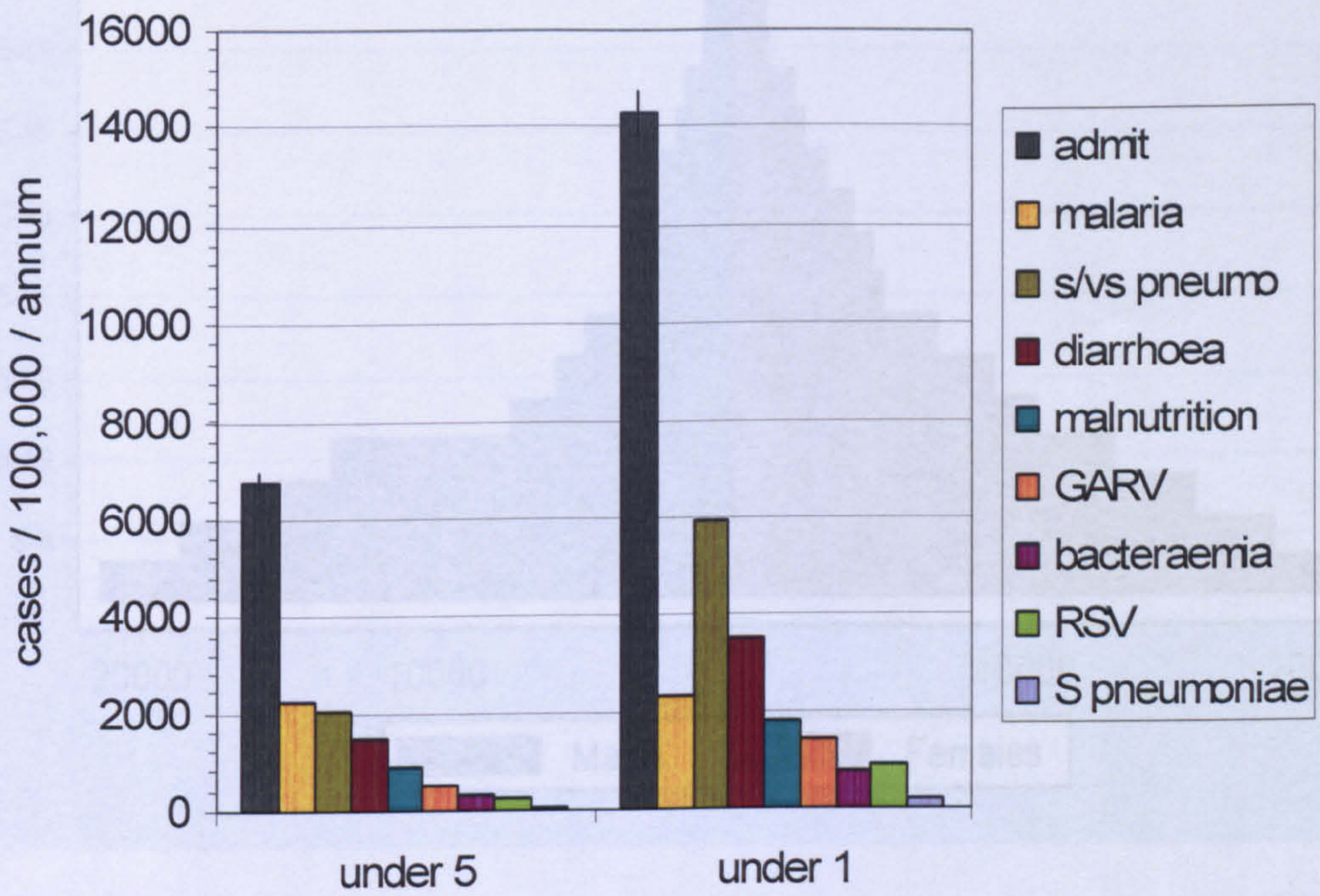


Figure 3.2. Comparative annual incidences of various diseases within the DSS over a five year period, 2002-05 by age group. s/vs- severe/ very severe pneumonia, GARV- Group A rotavirus.

3.2.3 The Population

The local people (within the DSS) are predominantly Giriama, one of the nine Mijikenda tribes of the East African coastal region. Whereas traditional cultural and religious practices like polygamy are still adhered to, other religions particularly Islam and Christianity have infiltrated the community. A substantial number of men from this region aged 20-50 are not resident within the DSS having migrated to urban areas such as Mombasa and Malindi

in search of employment mainly in the tourism and hotel industry as illustrated by the population pyramid (Figure 3.3).

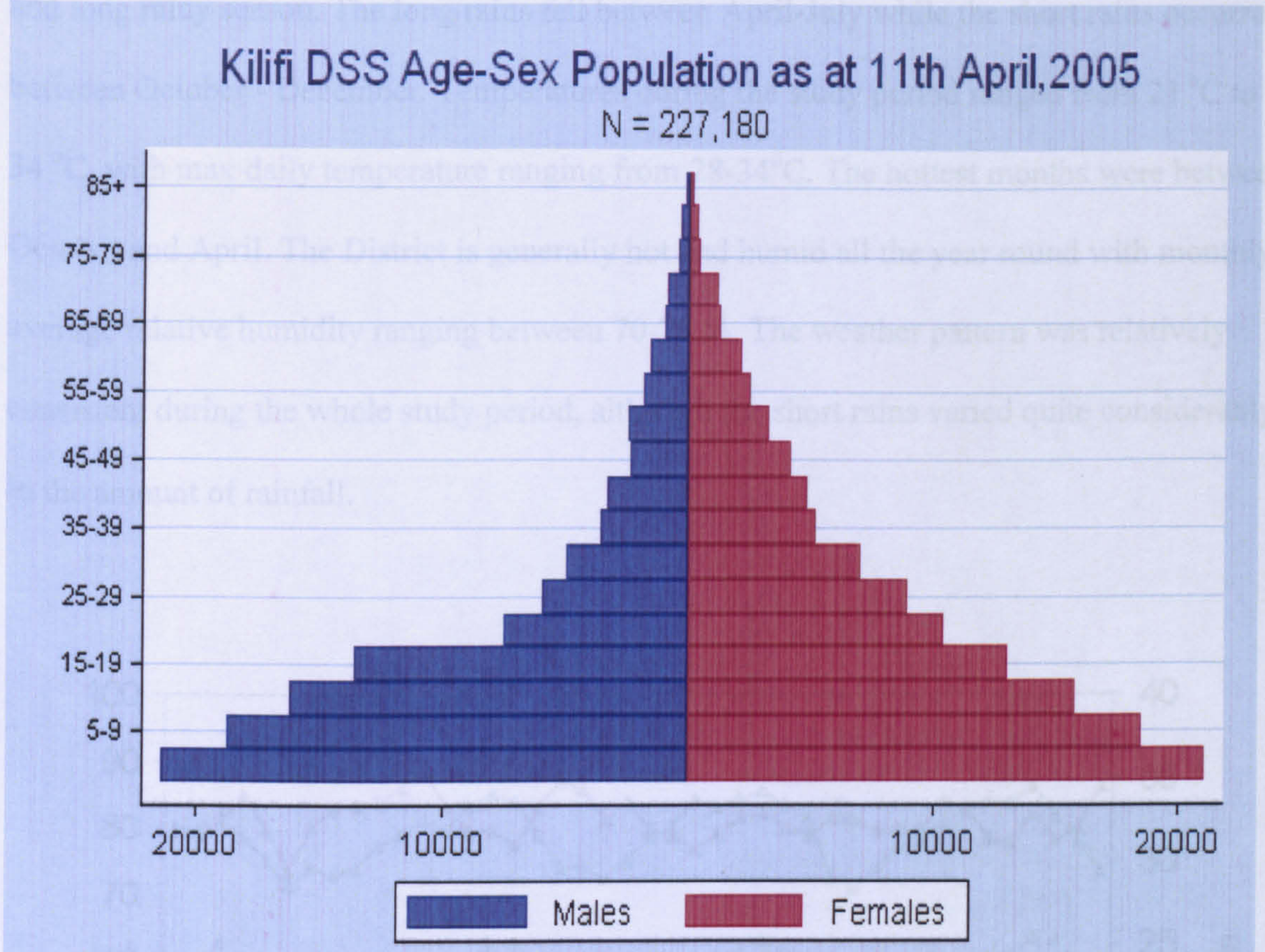


Figure 3.3. Population pyramid of the Demographic Surveillance System in Kilifi

3.2.4 Climate

Data used in this study was collected at the Kilifi Agricultural Institute located less than 2 kilometers away from Kilifi Hospital. Though not part of the national network of meteorological stations (the nearest one is 30 km away), data from the institute is routinely used by the research unit.

Daily data on rainfall, relative humidity and maximum and minimum temperatures was available for the whole follow up period (Figure 3.4). The climate of Kenya is tropical with the rainfall pattern in Kilifi district and Kenya as a whole divided into two seasons; short and long rainy season. The long rains fell between April-July while the short rains occurred between October - December. Temperatures during the study period ranged from 21 °C to 34 °C, with max daily temperature ranging from 28-34°C. The hottest months were between October and April. The District is generally hot and humid all the year round with monthly average relative humidity ranging between 70-78 %. The weather pattern was relatively consistent during the whole study period, although the short rains varied quite considerably in the amount of rainfall.

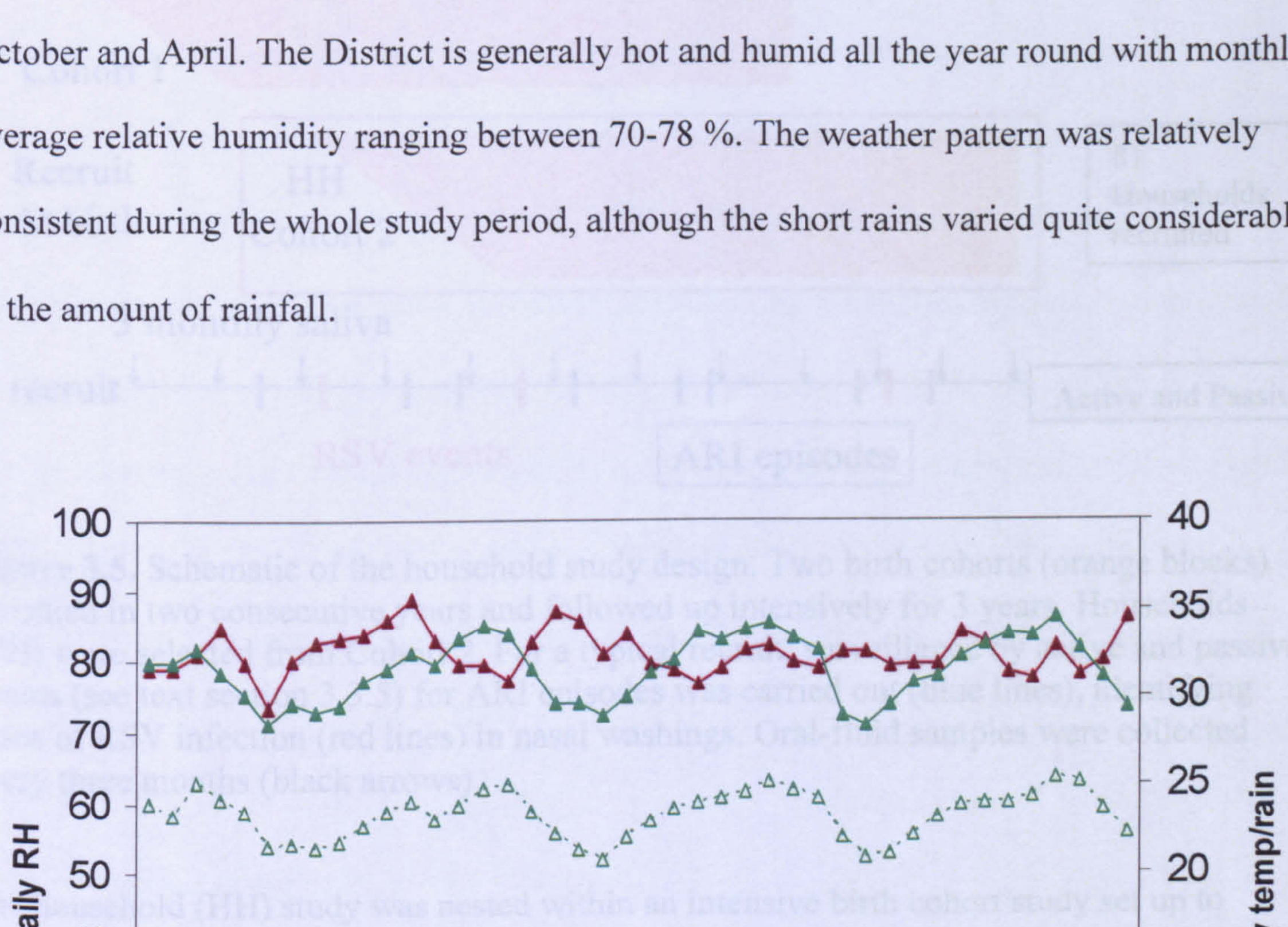


Figure 3.4. Climatic conditions in Kilifi district between the years 2002-2005. Shows for each month the average daily precipitation (mm), relative humidity (RH), and maximum and minimum temperatures.

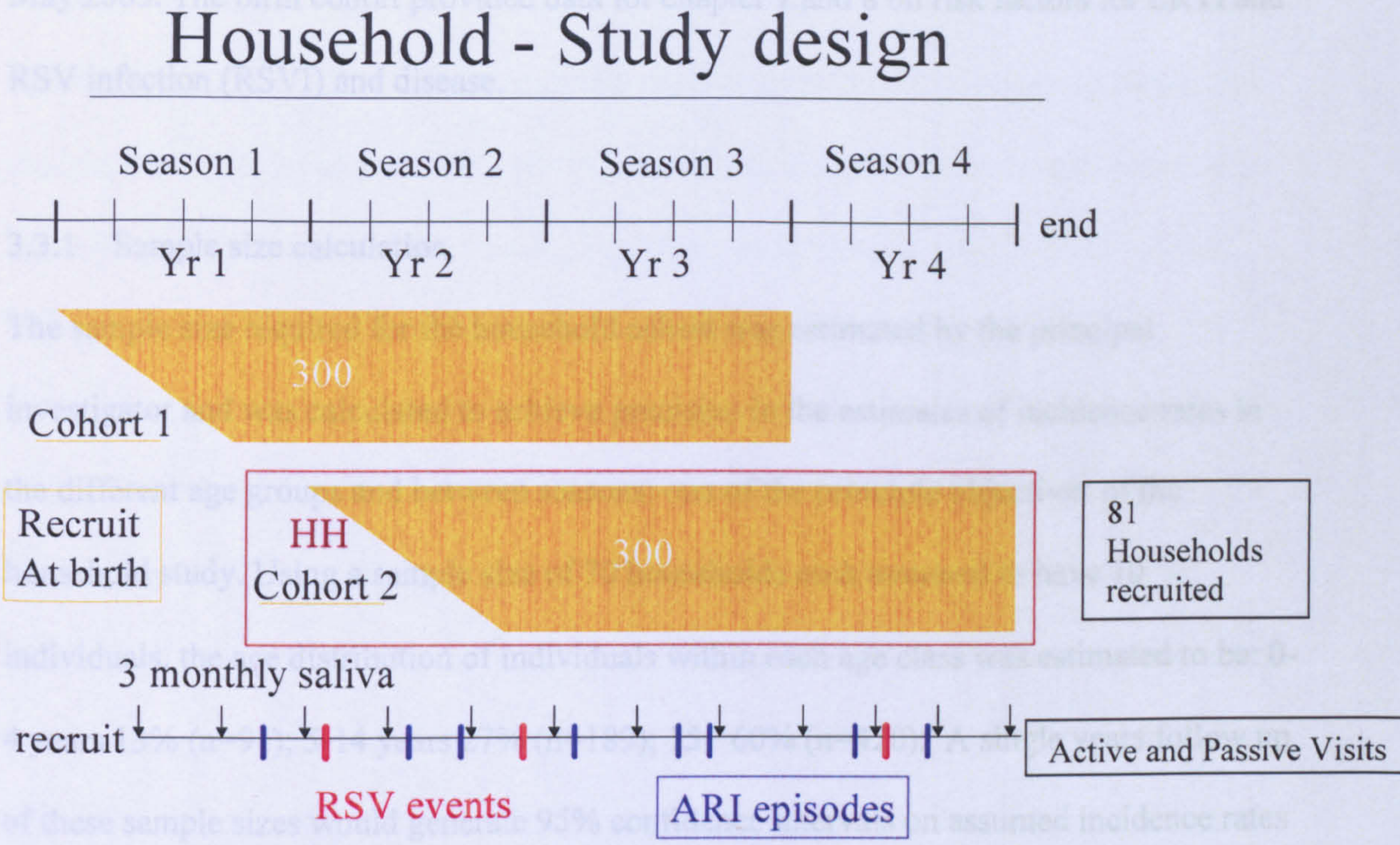


Figure 3.5. Schematic of the household study design. Two birth cohorts (orange blocks) recruited in two consecutive years and followed up intensively for 3 years. Households (HH) were selected from Cohort 2. For a typical recruit, surveillance by active and passive means (see text section 3.3.5) for ARI episodes was carried out (blue lines), identifying cases of RSV infection (red lines) in nasal washings. Oral-fluid samples were collected every three months (black arrows).

The household (HH) study was nested within an intensive birth cohort study set up to explore the natural history of RSV infection [5]. All occupants from a sub-sample of the families of these children formed the HH study population. Roughly, 600 infants were recruited at birth in the maternity ward at KDH or from the maternal child health (MCH) clinic if less than 2 weeks old. Recruitment was in two phases as shown on the study schematic: about 300 infants in the first year of the study and another 300 in the second year each followed for two to three calendar years. The first phase of recruitment was between January and May 2002 while the second phase was between December 2002 and

May 2003. The birth cohort provided data for chapter 7 and 8 on risk factors for LRTI and RSV infection (RSVI) and disease.

3.3.1 Sample size calculation

The sample size required for the household cohort was estimated by the principal investigator and was calculated to achieve precision in the estimates of incidence rates in the different age groups and between seasons: one of the principle objectives of the household study. Using a sample size of 70 households each assumed to have 10 individuals, the age distribution of individuals within each age class was estimated to be: 0-4 years 13% (n=91); 5-14 years 27% (n=189); 15+ 60% (n=420). A single years follow up of these sample sizes would generate 95% confidence intervals on assumed incidence rates [235-237] of 50%+/-15%; 15%+/-6% and 5%+/-2% /annum for each age group respectively. Considerably greater precision would be achieved over the whole study period.

3.3.2 Sampling design

Recruitment into the birth cohort was not intended to provide a representative sample of the community. The key aims of the birth cohort were related to immunity to infection rather than incidence estimation. Hence, for convenience recruitment was undertaken from deliveries at KDH, and (see next section) maternal child health clinic (MCHC). This sampling method afforded a relatively easy method of recruitment as opposed to random selection of births from within the community. Sampling was purposive insofar as to select within a defined region that would afford relatively easy access by field teams to the households (for active surveillance – see later section) and also by mothers/babies to the research OP clinic at the hospital (passive surveillance – see later section). The sampling

method also recognized the potential hazard of choosing all recruits from one small area, which might have resulted in (a) a complete absence of infection during an epidemic by chance, and (b) reduced variability of strains to which the study group was exposed. The method by which we established a sampling area was to conduct an exit poll from the MCHC clinic at KDH as described in the next section.

3.3.3 Selection of study area

Eight study locations from the DSS were selected as the study area. A location is an administrative demarcated area of the district (i.e. province-district-location-sub-location). Selection of these locations was done after an exit poll carried out at the KDH MCH clinic as mothers were leaving the clinic. The survey was undertaken primarily with the objective of (i) establishing if we could supplement recruitment from maternity with recruitment from MCHC, at an early age. (ii) Given that we were recruiting from the hospital we wanted to recruit those mothers/babies who lived within reasonable access of the hospital both to enable us to get to them, but also so that they were within reasonable access to the hospital (i.e. to facilitate OP clinic passive surveillance). An indication of this accessibility was to determine the locations from which mothers came to visit MCHC (rather than locations from which they came to give birth as this would be biased as some mums for example would travel very far if they had a problem birth).

150 mothers were randomly selected and a questionnaire administered enquiring on aspects of residence location, length of time to get to hospital and the travel costs incurred (Appendix A, p 287). The age of the child at first clinic visit was also determined as we wanted to assess the possibility of recruiting children from the MCH clinic. Results from this exit poll showed that majority of those presenting at the clinic came from 8 locations

namely: Ngerenya, Roka, Tezo, Kilifi Township, Chonyi, Matsangoni, Sokoke and Takaungu-Mavueni. These were selected as the study area.

3.3.4 Surveillance of infection

Surveillance of infections was by two methods, clinical and serological (Fig.3.5). Clinical surveillance comprised of two methods, that is, active and passive surveillance (section.3.3.6.), of respiratory infection. Active represents visits to households on a regular basis by field teams either weekly or monthly and passive represents visits to the OP clinic by study participants with respiratory symptoms. In each instance the identification of symptoms consistent with an acute respiratory infection (upper or lower) occasioned the collection of a nasal specimen by nasal wash method (section.3.3.7) and subsequent laboratory detection of RSV by antigen assay (section 3.3.9). Serological identification of infection was through repeated sampling of oral fluid every 3 months, by which to determine antibody boosting or seroconversion which would indicate infection (section.3.3.6 and Chapter 4). The reasons for the two methods of surveillance were: (i) clinical to identify actual time of RSV infection, determine genotype and severity, and hence things like WAIFW. (ii) serological to support estimation of age-specific incidence in the knowledge that clinical surveillance would under-estimate total infections (due to any or all of the following mild undetected or sub-clinical infection, cessation of shedding before sample collected, or sensitivity of antigen assay). A problem with serological identification of infection is that it would not identify with precision the time of infection and thus is not useful in determining sequences of infections.

3.3.5 Implementation

Field worker Training

The field workers (FW) involved in the project underwent extensive and intermittent training in various techniques. At the outset, they were trained on how to carry out the consenting process as well as the translation of the consent form into the local language in which they were all fluent. Training on obtaining consent was accomplished by means of role plays simulating the actual consent process. Role plays were also used to enact different case scenarios that could be encountered in the field. A second element of training was in the identification of clinical symptoms that were consistent with an RSV infection. This was achieved by showing the FW training videos as well as by the actual observation of sick children presenting to the ward with ARI under the supervision of a clinician. Thirdly, they received training on sample collection; collection of nasal wash samples and saliva samples and performance of venepuncture for blood collection. A field worker manual was developed with guidelines for both fieldwork and outpatient cover. It contained detailed accounts of procedures to be followed for different situations within the study. Training was ongoing throughout the study period with repeated training sessions every so often in order to maintain quality of performance.

Consent Form Development

Consent forms to be used in both studies were developed and submitted to the consent review committees (initial review locally by the Scientific Coordinating Committee of KEMRI-CGMR(C), and nationally at the Scientific Steering Committee of KEMRI and the National Ethical Review Board, Nairobi) for approval (Appendix B- E, p 289-301). This was followed by a translation of the forms from English into Kiswahili and Kigiriama, both of which are languages used by the local community. The process of translation and back

translation into English was done as a joint task between the project field workers who hail from the local community and spoke the language and the researcher in order to ensure that fieldworkers clearly understood the study and thus were able to translate accurately. The back translated version and the original version of the consent form were then cross-checked to ensure that no aspect of the original consent form had been omitted or its meaning altered during the translation process.

Recruitment of cohort children

A child was eligible for recruitment if (i) residing in one of the eight locations selected as the study area, (ii) one-way travel cost to their home was ≤50 Kenyan shillings, (iii) the journey time was <1 h . Some children were later excluded due to poor road access to their homes. Informed consent for participation in the study was obtained from each infant's mother or from both parents whenever possible before delivery or soon after birth. Children in this birth cohort were subject to two methods of surveillance, namely, active and passive. Active surveillance involved visits made to individual households by the study field workers and passive surveillance was by presentation to the KEMRI outpatient (OP) clinic encouraged for every episode of respiratory infection (including mild URTI e.g. including a runny nose). It is within this framework that we extended the second birth cohort study to include families of a selection of infants to be observed for the spread of RSV within the household.

Recruitment of Households

Recruitment of households was staggered over five months between February and May 2003. Household members of selected cohort infants (i.e. those with siblings) were asked to enroll into the study. After identification of eligible (Table 3.1) children from the larger

birth cohort, families of these children were visited soon after the delivery and discharge from the hospital to obtain consent to taking part in the household study. A family unit was defined as the individuals that normally eat and live together. This definition was used mainly because of the polygamous nature of members of this community and also because of the fact that extended family members tend to live together. The definition therefore covered the stepmother and stepsiblings of the birth cohort infant and in some cases aunts and cousins and meant that study families did not necessarily live in the same house but always within the same homestead. Figure 3.6 shows the spatial distribution of households in the study.

Table 3.1. The inclusion and any exclusion criteria for the households

Inclusion criteria	*Exclusion criteria*
(i) Born at the Maternity, or presents to the KDH CHC ≤14 days. Recruited form ward on primary infection	(i) consent with held
(ii) Resident within one of the 8 study locations	(ii) excludes family to children who died very young i.e. only those who survived for one visit contact (passive or active)
(iii) one-way travel cost to their home was ≤50 Kenyan shillings	
(iv) journey time to the hospital ≤1h	
(v) infants must have at least one other sibling	

Consenting Process

The consenting process for the main project took place at the hospital while that for the HH study took place at the family home. A deliberate attempt was made to have both parents present whenever possible. A trained fieldworker explained the details of the study in the local language and gave a translated copy of the consent form to the parents to read if they could. Information covered during the consenting process included details of the family's

involvement in the study including what would be required of them in terms of allowing visits to their homes by the field workers or visits by them to the research clinic; details on what, and how samples would be collected; anticipated risks and benefits of the study and issues of confidentially (Appendix C, p 293). Parents were informed that participation was entirely voluntary. If, on understanding their involvement the parents decided to take part in the study, they were asked to sign two copies of the consent form, one of which was given to them for reference purposes. Active and passive surveillance was initially limited to children i.e. any member of the family who was less than 15 years of age at the time of recruitment. At a later stage the protocol was revised to include adult family members (≥15 years old; see section below).

Inclusion of Adults

At HH enrollment all members were requested to participate in providing oral-fluid samples at 3 monthly intervals to identify infections by serological methods (see later section). These samples were to be screened for antibody changes to identify infection in adults and to calculate rates of infection. Initially, clinical surveillance was limited to children under 15 years of age. However, upon realization that the original study design (excluding adults from clinical surveillance) would not provide a clear picture of infections in the household, an amendment to the original proposal was submitted for review and approval received. Subsequently, at the beginning of 2004 all adults in the household study were asked to take part in routine clinical (active and passive) surveillance. This was done through a repeat consent process of individuals ≥15 years in study households (Appendix D, p 298).

Extension of Study

The initial study design planned to follow up children in the second cohort for two RSV epidemics only. However, in the epidemic of 2003 far fewer of the 300 infants in cohort 2 were identified as having been infected than was expected (see Fig 3.9b). As a consequence permission was sought and obtained from the review committees to extend follow up to 3 epidemics and therefore to monitor children (and families) for another RSV season that would extend the study into 2005. Thus in the last quarter of 2004 participants were asked to continue taking part in the study through to the end of the expected epidemic of 2005 (Appendix E, p 301).

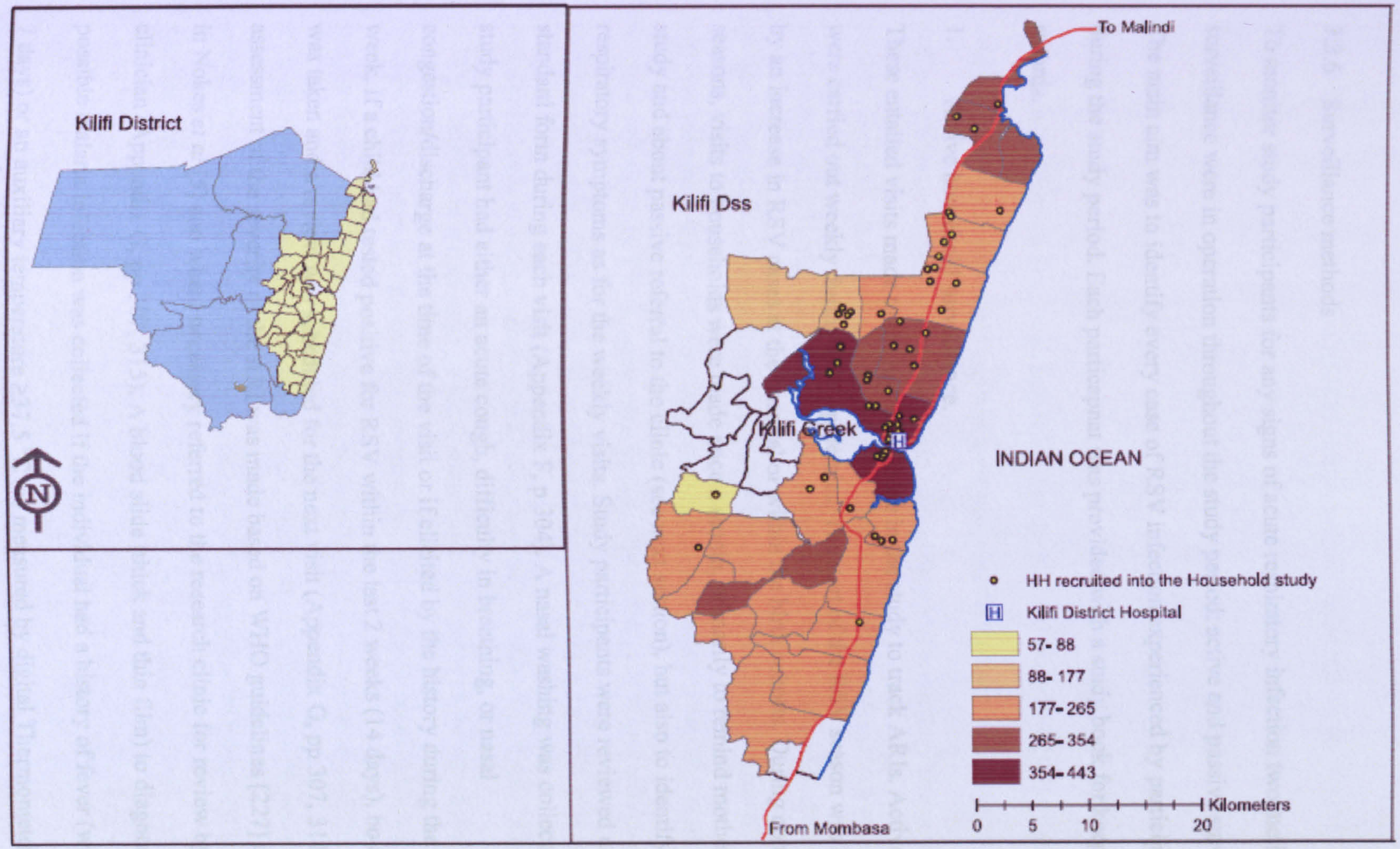


Figure 3.6. Map showing spatial distribution of households within the DSS on a background showing the population density of the area from which the birth cohort was recruited. The white areas are not part of the study area.

3.3.6 Surveillance methods

To monitor study participants for any signs of acute respiratory infection two methods of surveillance were in operation throughout the study period: active and passive surveillance. The main aim was to identify every case of RSV infection experienced by participants during the study period. Each particapnat was provided with a study book for keeping visit records.

1. Active household surveillance.

These entailed visits made to the homes of those in the study to track ARIs. Active visits were carried out weekly during RSV epidemics. The start of the RSV season was identified by an increase in RSV cases at the hospital or within the birth cohorts. During non-RSV seasons, visits to households were made once a month, primarily to remind mothers of the study and about passive referral to the clinic (see next section), but also to identify respiratory symptoms as for the weekly visits. Study participants were reviewed using a standard form during each visit (Appendix F, p 304). A nasal washing was collected if a study participant had either an acute cough, difficulty in breathing, or nasal congestion/discharge at the time of the visit or if elicited by the history during the past week. If a child had tested positive for RSV within the last 2 weeks (14 days), no action was taken and a review was scheduled for the next visit (Appendix G, pp 307, 314). An assessment of the severity of the ARI was made based on WHO guidelines [227] described in Nokes *et al* [5] and where necessary referred to the research clinic for review by a clinician (Appendix G, pp 307, 315). A blood slide (thick and thin film) to diagnose a possible malaria infection was collected if the individual had a history of fever (within prior 7 days) or an auxiliary temperature ≥37.5 °C (as measured by digital Thermometer, Becton-Dickinson, USA) on the day of the visit. Participants with temperature ≥37.5 °C on the day

of the visit were referred to the KEMRI research clinic (Appendix G, pp 307, 313). A record of the number of complete weeks in between active hospital visits was kept to facilitate an accurate record of the days of observation.

2. Passive surveillance

Passive surveillance was through self-referral to the KEMRI outpatient (OP) research clinic located in KDH. All year round mothers were encouraged to present to KDH (i.e. self/parental referral) with an infant or child who had signs of respiratory disease that was possibly caused by RSV i.e. ranging from mild URTI including a runny nose or difficulty in breathing to LRTI. Parents of infants and children in the study were primed on how to recognize symptoms consistent with a possible RSV infection. The clinic operated on weekdays between 8 a.m. – 5 p.m. Study participants presenting at the OP research clinic were interviewed by a FW to assess need for a nasal washing, and to take measurements (oxygen saturation, axillary temperature, weight and heart rate) and subsequently reviewed by a clinical officer using a proforma method to ascertain the severity of ARI (Appendix H, p 323). The criteria for collecting a nasal washing were the same as for active visits. Transport costs were reimbursed and treatment provided without charge. Limitations of support for other members of the family attending the OP clinic operated as defined in the form appended (Appendix I, p 328). Children referred to the clinic by a FW undertaking an active visit to the house, would be reviewed by a CO as described above using the proforma detailed in Appendix H.

3.3.7 Study processes

1. Follow-up visits

Follow up visits for a positive slide or nasal washing are as detailed in Appendix G (pp 316-317). In the case of absent participants, an inquiry was made with the next door neighbours on the whereabouts of the study participants. If they were said to be close by then the FW either waited at the homestead for the participants or tracked them. Follow up of unavailable participants was done either the next day or at the weekend.

2. Sample collection and processing

Nasal specimen

Nasal washings were collected by use of an ear syringe bulb (Abbort, Ashland, Ohio-USA) using a method described in the standard operating procedure (SOP) in Appendix G, p 319. The method is based on that described by Hall and Douglas [238] with minor modifications [5]; essentially normal saline (quantity according to age) is squirted into one nostril and collected from the other (Fig. 3.7). Samples collected in the field were stored in cold boxes and brought back to the microbiology laboratory at the research centre at the end of each field day where they were stored in a -4^{o}C freezer. Those collected in the OP clinic were delivered within the hour. Nasal washing samples were processed usually within 24 hours or invariably within 4 days usually over the weekend. First, any mucus in the sample was broken up, and then slides were prepared from 200 ml of specimen by use of a cytology centrifuge (Cytospin 3; Thermo Shandon Ltd, Cheshire UK; 67 *g* for 10 min) (Appendix J, p 330) and screened for RSV antigen by use of a direct immunofluorescent antibody test (DFA) (Light Diagnostics RSV screen; Chemicon, Temecula, CA USA), according to the manufacturer's protocol described later. Samples were assigned as either positive or negative for RSV antigen, on the basis of DFA result.

Figure 3.7. Nasal specimen collection by nasal wash (NW) bulb method in the clinic and in the field setting

Oral-fluid samples

Oral fluid was collected using a sponge swab (Fig. 3.8) known as Oracol (Malvern Medical Developments, Worcester, UK). This sponge swab consists of a cylinder of expanded polystyrene foam attached to a plastic stick and is used like a toothbrush [212]. The swab was brushed along the gums and mouth for 60 seconds and the device then inserted into a plastic tube, stoppered, stored on ice and returned to the laboratory upon return from the field (Appendix G, p 320) or within one hour when taken in the OP clinic. In the laboratory 1ml of preservative buffer (20% foetal calf serum and 0.2% sodium azide in PBS) was added to the sample. The oral fluid was then squeezed out of the foam and clarified by centrifugation and the supernatant stored in a -70 °C freezer for processing at a later stage using both the IgG Capture Assay and the direct ELISA (described in Chapter 4).

Figure 3.8. Oral-fluid collection in the field using an Oracol device

Blood samples

Acute and convalescent samples were collected by venepunture (Appendix G, p 321) at each episode of RSV infection identified by IFAT on NWs. The acute sample was collected as soon as possible after a positive IFAT result was received from the laboratory and a convalescent sample collected one month later. Once in the laboratory, the clotted blood was centrifuged and serum stored for processing by direct ELISA at a later stage (Appendix N, p 344).

3.3.8 Quality control (reliability)

1. Quality Control for collection of data

Each FW was provided with a field manual containing all study details. They had the manual at hand at all times enabling referral for clarification of any study details. It was impressed upon the FW to ensure forms and questionnaires were completed carefully and as required. They were instructed not to make assumptions about answers elicited but rather to ask the question again if they were unsure. The FWs were trained to be methodical in their actions as well as to make sure they were clear on exactly what they intended to do before arriving at the household. At the end of each visit, a review of the entry was carried out before leaving the home. On their return to the unit the FW had to (i) go through their individual field forms to ensure that no blatant mistakes had been made (ii) relate any problems encountered in the field to the field co-ordinator (the researcher) and also note them down for discussion during fieldworker meetings and, (iii) enter the rsv number of individuals from whom nasal washings had been collected on each day into the sample book in the OP clinic. This was updated as and when the NW result was received from the laboratory and any follow-up appointments were made as necessary. Active visit data sheets were checked at the end of each field day either by the field work coordinator or by a senior project FW for any inconsistency and were then given to the data entry clerk.

As an additional monitoring system to ensure quality of data and samples collected, senior project members accompanied the fieldworkers on the household visits during certain days of the week on a rotational basis. The aim was to ensure that study questions were asked correctly, that samples were collected in the proper manner and that the information provided by the mother was entered into the forms correctly. A similar process was also carried out in the outpatient clinic. Any mistakes made were reviewed and the fieldworker

advised accordingly. Training of the field workers on study protocol and sample collection was done periodically throughout the whole study period.

2. Quality Control for sample collection

Strict measures were taken to ensure that i) all bottles and tubes for saliva, nasal specimen and blood collection were labelled before any sample was taken ii) all samples had the individual's name, serial number, sample designation and the date of collection of the sample iii) saliva samples were collected for one minute to ensure uniformity (Appendix G, p 320) and iv) that the saline solution used in the nasal wash reached the nasopharynx (i.e. observing saline emitted from the alternative nostril to that in which it was instilled). If this was not verified the procedure was repeated (Appendix G, p 319). Every effort was then made to return collected specimens to the laboratory as quickly as possible, on ice as required. For specimens that were spoilt or inadequate (e.g. slides or nasal washings), or where specimen tubes could not be read a fresh specimen was requested.

3.3.9 Laboratory Methods

1. Direct Immunofluorescent Antibody Test (IFAT).

An immuno-fluorescent antibody test (Light Diagnostics RSV DFA, Chemicon, Temecula, CA USA) was used to detect RSV antigen in the nasal washing collected. The test utilizes a direct immunofluorescent antibody technique for identifying RSV in cell preparations made from nasopharyngeal aspirates (NPA) or nasal washings. The monoclonal antibodies to the viral antigen are labelled with FITC (fluorescein isothiocyanate), which fluoresces apple-green when illuminated with ultraviolet light. The labelled antibody binds to the viral antigen present in the specimen and any unbound reagent is removed by washing with buffer. Cells in positive specimens will fluoresce apple–green, while uninfected cells stain

dull red due to the presence of the counter-stain Evans blue (IFAT SOP – Appendix K, p 331). The IFAT assay was performed by the project technicians (John Abwao and Anne Bett).

2. Molecular methods for genotyping RSV isolates

RSV positive cases identified by antigen IFAT were subjected to molecular characterization. The method involved RNA extraction and subsequent by reverse-transcription (RT) to obtain complementary DNA (cDNA), followed by a multiplex PCR to identify RSV or influenza target sequences within the sample. Multiplex PCR uses a combination of several primer pairs in the same amplification reaction, with the objective of producing different specific amplicons depending on the target present in the sample. This method is increasingly used for the diagnosis of infectious diseases caused by DNA/RNA containing viruses in which the starting nucleic acid template may be low in copy number or of poor quality compared to chromosomal DNA [239]. In this case the method was set up in the laboratory for general respiratory viral diagnostic reasons and adopted in the present study. Further characterization of MPX positive samples involved genotyping of the virus based on restriction fragment length polymorphism (RFLP) analysis of the N or G genes and selective sample sequencing of the G gene (see reference Scott *et al.* [83].

(i) RNA extraction and cDNA synthesis

RNA was extracted from RSV positive samples using QiAmp Viral RNA kits (Qiagen UK, Ltd), according to the manufacturer's instructions (supplied with the kit). The addition of the carrier RNA to buffer AVL is not done. cDNA was synthesised using the Omniscript RT kit (Qiagen UK Ltd). Briefly, 4 µl 10X RT buffer, 4 µl 5 mM dNTPs, 250 ng pdN(6) random hexamers (Amersham Pharmacia, UK), 1 µl Omniscript reverse transcriptase, 10 µl

nuclease free water and 20 µl RNA were mixed and incubated at 37°C for 1 hour. The cDNA was either used directly or stored at -20 °C for later use.

(ii) PCR and N and G gene RFLP

All positive samples were subjected to an RSV/influenza multiplex PCR (MPX) based on the method developed by Stockton *et al* [239] as detailed in the SOP (Appendix L, p 336). N Gene PCR was carried out on the first round products from the RSV/Flu multiplex assay. RSV N gene typing was done according to procedure described by Cane *et al* [240] and RSV G gene typing by RFLP as described in Scott *et al* [83] and detailed in the SOP in Appendix L (p 336). These molecular assays were performed by Anne Bett one of the project technicians and Dr. Paul Scott (Post doc on the project).

3.3.10 Data Handling

Data collected in the field was entered the next day into a FileMaker Pro 5.5v1 Dev, (FileMaker, Inc. Santa Clara, CA USA) RSV study database. The entered data was then checked by second data clerk to ensure correct entry of study details. Any inconsistencies were corrected. Data collected in the OP clinic was entered directly onto the study database on a real time basis (as for all data at admission review). The database had inbuilt consistency and range checks to minimise errors of entry (although cant ensure they are correct). All the data was analyzed using Excel (Microsoft Office XP), and STATA software Version 8.0 (Stata Corp., College Station TX, USA). Details on the analysis conducted are described in Chapters 5-9.

3.3.11 Ethical considerations

The study underwent local internal review by the unit Scientific Coordinating Committee prior to submission to KEMRI Scientific Steering Committee. Subsequent ethical approval for the study was sought and granted by the Kenya Medical Research Institute/National Ethical Review Committee and Coventry Research Ethics Committee. Any amendments to study procedure went through the same ethical approval process.

3.4 General Results

3.4.1 Recruitment

A total of 81 households were recruited into the HH study, being associated with 84 birth cohort children (there were twin birth cohort children in three instances.) The vast majority (79) were selected from birth cohort children recruited at delivery in the Maternity Ward. Of the 73 households under surveillance in 2004, the adults members of 41 households agreed to participate in clinical surveillance and the remaining 32 refused. The main reason given by adults in these household was- they did not want to have a nasal wash. No differences were identified between these HHs and those that agreed to participate. Households lost to follow up before the second epidemic (8) were replaced with HHs of other birth cohort children (Figure 3.10). Replacement was with those households in which the birth cohort child had not been identified (by clinical surveillance) as having already been infected. The 81 households had a total of 447 individuals, 133 (29.8%) of whom were adults (≥15 years) and 314 children; 84 (18.8%) birth cohort children and 230 (51.5%) were siblings less than 15 years of age at the time of recruitment. Specific details of households recruited and characteristics of study participants have been provided in Chapter 5.

3.4.2 Loss to follow-up

In the main cohort 31% of the children were lost to follow-up. The risk of loss to follow-up in the household cohort was the same 31% (25/81) households. 16% (13) of the families moved out of the study area while 14% (11) withdrew consent before the end of the study. One household was removed from the study as a result of the death of the birth cohort child as recorded in Table 3.2. 36% (4) of the households that withdrew consent did so at the time when follow-up was being extended (i.e. late 2004).

Table 3.2. Reasons for exiting the HH study

Reason for exiting the study	Frequency	Percentage
Died	1	1
Moved	13	16
Refused	11	14
Study end	56	69
Total	**81**	**100**

To assess whether the households that were lost to follow up were a biased group several characteristic were assessed: household size, number of pre-school and school children, travel time and travel costs to hospital, respiratory episodes and social economic status. There was no evidence to suggest an association between any of the factors and drop out (χ^2 test, $P>0.05$) except in the case of number of pre-school children. The proportion of pre-school children in households that remained in the study was significantly higher than in households that were lost to follow up ($\chi^2(1) = 6.151$, $P = 0.013$).

3.4.3 Epidemics in the main birth cohort and in individuals of the household cohort

The working definition of an epidemic was a period delimited by weeks in which greater than or equal to one case was found and within which at least three cases were found in any

contiguous 3-week period [5]. The interval between epidemics (inter-epidemic period) was defined as the time interval between the end of one epidemic and the start of the next epidemic. Between the years of 2002 and 2005 during which this project was carried out there were 4 RSV epidemics within the dates 12/3/02 to 1/7/02, 3/12/02 to 15/4/03, 8/1/04 to 2/6/04 and 11/11/04 to 18/2/05. The first cohort experienced three epidemics while the second cohort recruited towards the end of 2002 extending into 2003 experienced the tail end effects of the second epidemic as well as epidemics three and four as shown in Figure 3.9.

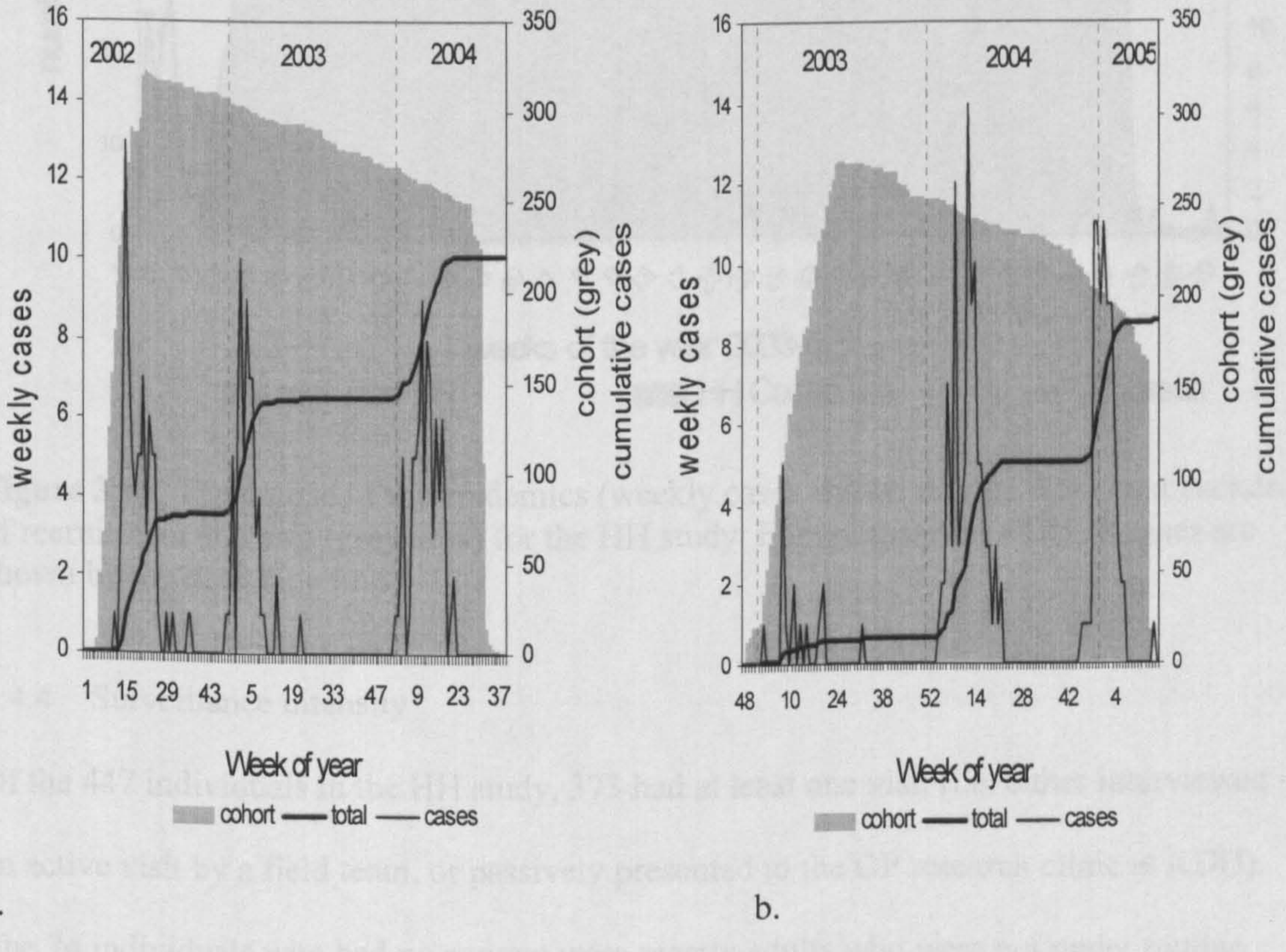


Figure 3.9. Temporal pattern of RSV infection in the birth cohorts: panel (a) cohort 1, panel (b) cohort 2. Grey area shows cohort size at each time point, and ascending line from left to right in each panel shows the accumulation of the shown weekly cases.

Households that were lost to follow up before the epidemic of 2004 were replaced with other households for which there was no evidence (from clinical surveillance) that the

index child had been infected in the 2003 epidemic. After this initial replacement of households, no further replacement was undertaken (Figure 3.10).

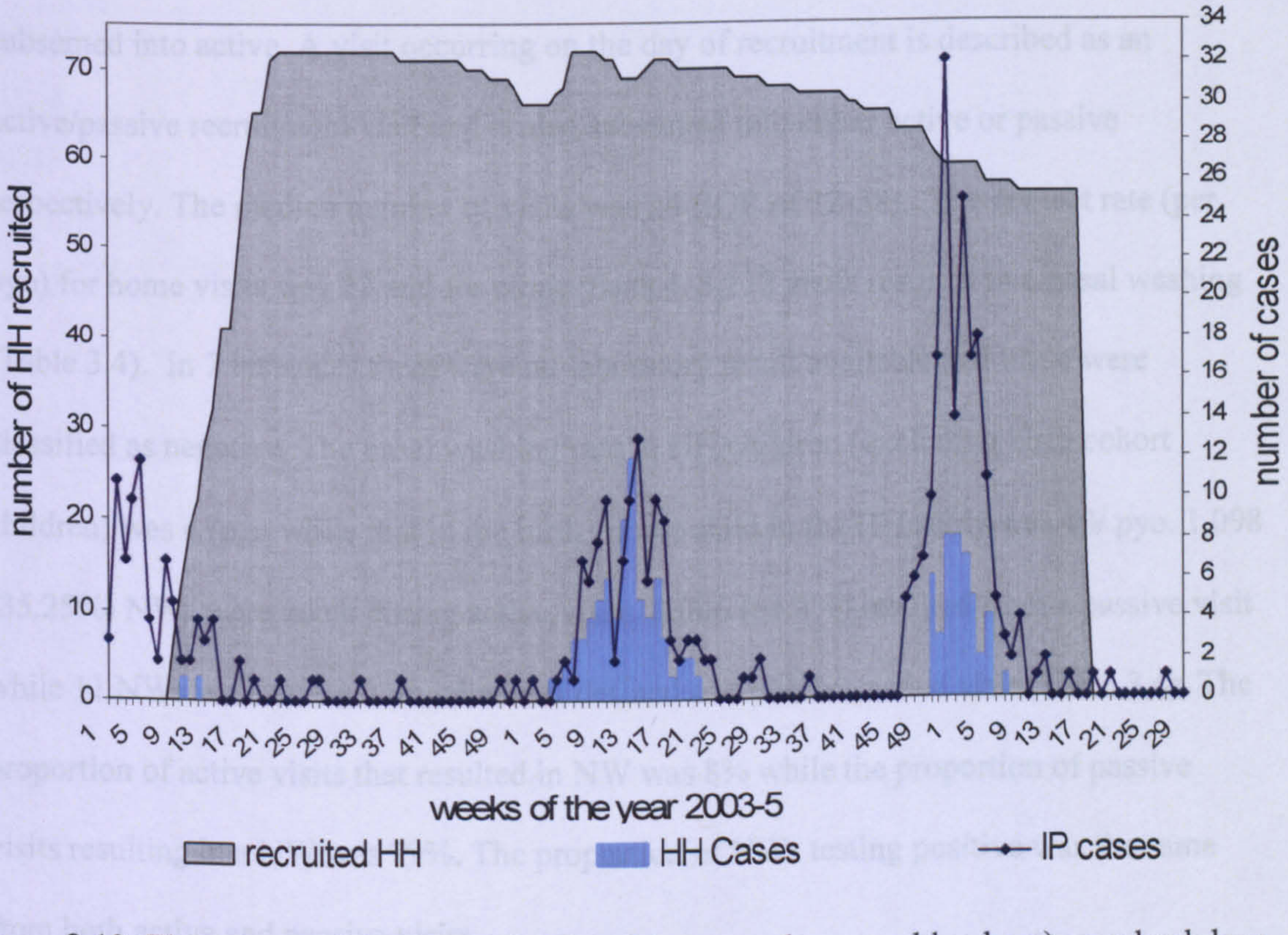


Figure 3.10. The course of the epidemics (weekly cases shown as blue bars) on a backdrop of recruitment and exit (grey area) for the HH study. For comparison KDH IP cases are shown by the dark blue line.

3.4.4 Surveillance intensity

Of the 447 individuals in the HH study, 373 had at least one visit (i.e. either interviewed in an active visit by a field team, or passively presented to the OP research clinic at KDH). The 74 individuals who had no contact were mostly adults who were not under routine surveillance (only part of the serological – oral fluid – surveillance.) In total, follow-up amounted to 626 person years of observation (pyo),which represents the summation of time for each individual from recruitment to study end or drop out, excluding periods spent away from the DSS (see Active visit form Appendix F, p 304). A total of 16,115 visits (active/passive) were made (Table 3.3). A passive-active visit was an active visit that

resulted in a referral to the research clinic by the FW due to the severity of infection as discussed previously. Passive visits arising from active of which there were 15 were subsumed into active. A visit occurring on the day of recruitment is described as an active/passive recruitment visit and is also subsumed into either active or passive respectively. The median number of visits was 24 (IQR of 12-38). The contact rate (per pyo) for home visits was 22 and for clinic visits 4. 3,122 visits resulted in a nasal washing (Table 3.4). In 7 instances there were no laboratory result available and these were classified as negative. The nasal washing rate in HH children (excluding birth cohort children) was 4 /pyo while that in the birth cohort child in the HH study was 10/ pyo. 1,098 (35.25%) NWs were taken during active visits, 2,006 (64.4%) resulted from a passive visit while 11 NWs were taken from what was defined as a passive-active visit (Table 3.4). The proportion of active visits that resulted in NW was 8% while the proportion of passive visits resulting in a NW was 79%. The proportion of NWs testing positive was the same from both active and passive visits.

Table 3.3. Details on surveillance intensity

	Cohort type			
visit type	**adult**	**birth cohort**	**child**	**Total**
active	1,694	3,295	8,574	13,563
passive	151	1,256	1,145	2,552
Total	**1,845**	**4,551**	**9,719**	**16,115**

Table 3.4. Details on nasal washings taken during the study

	Nasal wash results (%)		
visit type	**negative**	**positive**	**Total**
active	1,062*	48 (4.3)	1,110
passive	1,927*	85 (4.2)	2,012
Total	**2,989**	**133 (4.4)**	**3,122**

* 7 children had no results and are assumed to be negative

3.4.5 RSV infection

Virus was introduced into 65% (53) of households during the whole follow-up period. There were 121 separate RSV related clinical infections – 1 adult was positive, and there were 56 infections in the siblings and 64 infections in the birth cohort child (Table 3.5). 79 infections were identified from a passive visit while 42 infections were from an active visit.

Table 3.5. RSV infection in the household study

	Cohort type			
Number of RSV infections	**adult**	**birth**	**sibling**	**Total**
1	1	49	49	99
2	0	13	6	19
3	0	2	1	3
Total	**1**	**64**	**56**	**121**

Definitions: adult is ≥15 years, sibling is brother or sister to a birth cohort child (=birth) and <15 years old

3.4.6 Seasonality

The epidemics observed during follow-up did not appear to be associated with the rainy season; epidemics occurred in the hottest months with lowest rainfall (Figure 3.11).

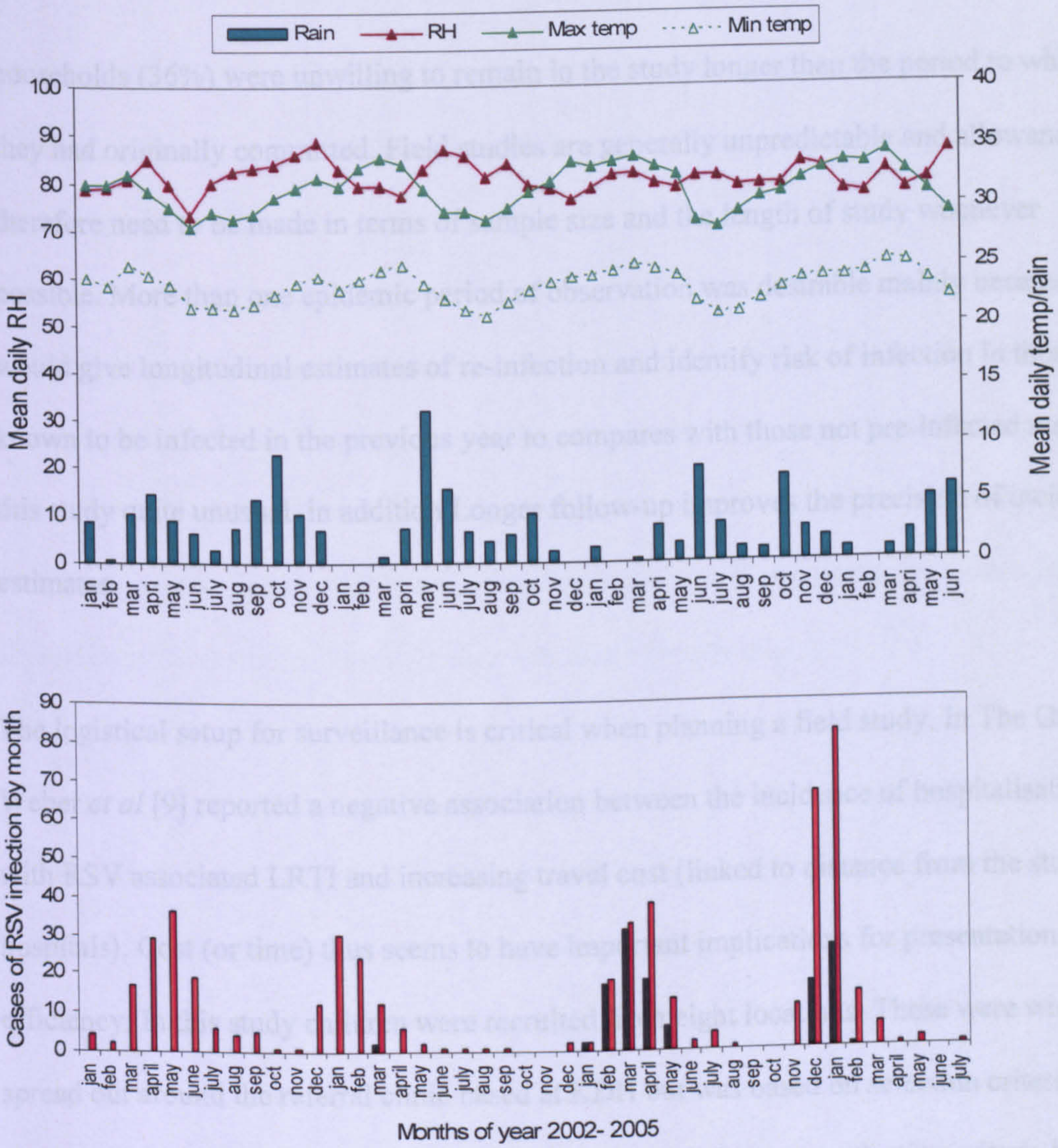


Figure 3.11. Cases of infection in the Household cohort (black bars) and in the IP (red bars) correlated to meteorological data during follow-up.

3.5 Discussion

The household study was scheduled to run for 2 years or two RSV seasons (that is until the end of 2004). However, this was extended to a third epidemic (that is until 2005) because the epidemic of 2003 began around 3 months earlier than expected (i.e. around 9 months after the preceding epidemic) and the cohort was scarcely exposed during the first year of follow up (i.e. there were very few infections identified). This could explain the increase in rate of drop outs at the end of year 2, due to refusal to continue, as some of these

households (36%) were unwilling to remain in the study longer than the period to which they had originally committed. Field studies are generally unpredictable and allowances therefore need to be made in terms of sample size and the length of study whenever possible. More than one epidemic period of observation was desirable mainly because it would give longitudinal estimates of re-infection and identify risk of infection in those known to be infected in the previous year to compares with those not pre-infected making this study quite unusual. in addition Longer follow-up improves the precision of incidence estimates.

The logistical setup for surveillance is critical when planning a field study. In The Gambia, Weber *et al* [9] reported a negative association between the incidence of hospitalisation with RSV associated LRTI and increasing travel cost (linked to distance from the study hospitals). Cost (or time) thus seems to have important implications for presentation efficiency. In this study children were recruited from eight locations. These were widely spread out around the referral clinic based at KDH but was based on selection criteria used to facilitate relatively easy access for passive referrals. Moreover, selection criteria based on accessibility are also of importance for active surveillance by field teams and it may be necessary to limit the area from where study participants are recruited to maximize successful contacts, or make manageable the number of household visits required. There is a need to remember that a scheduled visit is not always successful necessitating repeat visits and hence requires an additional allowance of time. Other issues to consider are: i) accessibility to homes (i.e. the road conditions or whether one exists) ii) available modes of transportation and a contingency vehicle is essential. The use of a motorbike for follow-up was found to be a better alternative to a vehicle where road accessibility was particularly problematic and in remote parts with few households to visit where a motorcycle was

considerably more cost-effective. However, special training is often required and road accidents (none more serious than a broken leg in this study) a threat to the smooth operating of surveillance.

Balancing the requirements of ease of access and a manageable study area for field visit is the recognition of chance events in infectious disease transmission. The recruitment area was the largest manageable within budget constraints in order not to suffer the possibility that the study population might miss an epidemic and that it would not experience a limited repertoire of viral strains if each variant was spatially restricted.

Previous studies reviewed by Nokes [38] suggest that a combination of methods of surveillance is required to identify all cases of RSV hence in this study we adopted both clinical and serological surveillance. Crucial to this study however, was not only to identify if someone had been infected but the time of infection by which to infer from whom infected and (for the main project) by what strain. Hence in an effort, to improve ascertainment by clinical means we adopted both passive and active surveillance. A majority (65%) of the ARIs were identified through a passive visit implying that study participants were conscientious about presenting to the OP clinic. For the typical passive visit there was a high probability that the symptoms would indicate the need to collect a nasal washing (79%), compared to that for an active visit (8% - χ^2, $P< 0.001$). This suggests that parents well understood the reasons given for bringing their children to the OP clinic. Interestingly, the contribution of RSV to acute ARI in the children was the same whether identified by a passive or active visit. Thus both proved very useful. Even so, it is likely that many infections were probably missed as there are significant problems in the clinical surveillance: mild or sub-clinical cases, end of shedding before sampling, low sensitivity of

IFAT if shedding is low [125, 241]. And unfortunately the oral fluid method which was intended to compliment clinical surveillance failed (reviewed in Chapter 4).

Our surveillance methods may possibly have been improved upon if the study design had been such that: i) households were visited every 2-3 days, ii) nasal washing were taken even when individuals did not displaying signs of respiratory infection, and iii) used a more sensitive method of virus detection as in PCR for all samples. This would have been logistically near to impossible, maybe difficult to justify ethically (i.e. since the nasal wash is not without some pain) and also very expensive. There are problems also in terms of method of taking nasal sample i.e. washing in adults which proved tricky.

Studies that include older children and adult participants must give careful consideration to sample collection. In this study collecting a sample from adult participants was a challenge as most adults were not keen on having a nasal specimen taken, viewing it as a strange and uncomfortable procedure. The bulb also had to be modified for use in adults; needing to be cut shorter to prevent penetration too deep into the nostril which can cause additional discomfort. The original study proposal did not have approval to collect nasal samples from adults and therefore this had to be sought at a later stage. The fact that adults were initially not under routine surveillance and were only included later in the study, resulted in incomplete data. Future studies would need to evaluate the acceptability of sampling procedures in each participant group prior to the start of the study.

The method of collecting nasal specimens adopted in this study was the Nasal Washing. Although Nasal Pharyngeal Aspirates (NPAs) are regarded as the most appropriate specimen from which to recover pathogens that do not colonize the nasal pathway [242],

NW is considered to be the optimal specimen for rapid antigen tests [125] and is also feasible for use in larger field studies. A study carried out in Kilifi [243] showed that samples collected by NW were of good quality (82% had >20 epithelial cells). Titres of RSV virus in NW specimens have been shown to be similar to those in NPA samples, and markedly higher than in nasopharyngeal swabs [238, 244, 245]. The applicability of NPA outside the clinic setting is questionable. It requires a mechanical or electric suction device and with apparent greater invasiveness is likely to create greater anxiety to children and their parents. The nasal washing method has previously been used in community based studies [94, 246] and the experience in the present study was that it was generally well accepted within the family for infants and children – as much or more by the parents as by the children themselves. However, as described above use of the NW on adults was not well received.

Diagnosis of RSV infection can be made by several methods; virus isolation, detection of viral antigens, detection of viral RNA, demonstration of a rise in serum antibodies, or a combination of these approaches. The present study adopted the direct immunofluorescent test (IFAT). Early studies have used the more specific culture methods for virus isolation although most clinical laboratories now use antigen detection assays to diagnose infection. Studies that have compared different virus detection methods have shown immunoflourescence techniques to be equally or more sensitive for the detection of RSV infection than culture [247-250]. The review by Weber *et al* [28] clearly showed that following the introduction of IFAT in studies, the proportion of ARI patients in which RSV was identifiable increased considerably (Fig 2.1). However, the ideal is a combination of culture, IFAT and perhaps PCR where possible [248].

Of importance is the quality of the samples collected. For a reliable IFAT, the integrity of the specimen must be maintained as it requires an adequate number of epithelial cells for acceptable sensitivity [125, 251]. Falsey *et al* [241] found that the direct antigen tests often lacked sensitivity for specimens obtained from adults and older children which may explain why there was only one positive sample identified from an adult in this study. It is thought that the reduced performance of the test may be due to a combination of certain factors in adults: lower viral titres, a shorter duration of shedding and dry mucosa [125].

One of the advantages of using the antigen detection assay in this study was that the results were available within a relatively short time enabling the collection of an acute sample, an important component of the project. Also this study was nested within the main birth cohort and immunofluorescence is the recommended RSV diagnostic assay for samples from children. To investigate some of the sensitivity problems we may have with the diagnostic method in this study, it is possible to screen all the nasal specimens that were negative by IFAT test using Polymerase Chain Reaction (PCR) based methods to identify any false negatives. PCR offers enhanced sensitivity combined with rapid diagnosis and sub-typing ability [239]. Until recently, PCR for pathogens with RNA genomes has been limited because of the difficulties involved in the RT (reverse transcription) step of the method and in the extraction of the nucleic acid when the starting material is of poor quality [239]. It is also relatively expensive compared to IFAT. Rapid ELISA was an option, but its sensitivity is lower than for IFAT.

Training procedures, methods of assuring quality of performance and related materials used at the KEMRI-WT operated paediatric wards and out-patient research clinic were established as a result of the studies carried out by English *et al* [252, 253]. These

procedures were in place prior to the start of the present study, and totally independent. All clinical officers (CO- clinical workers with 4 years of medical training) in the present study had previously undergone this training process. Respiratory rate was assessed by direct observation of the number of breaths taken per 30 seconds and scaled up to one minute. Clinical assessment was therefore based on previously established and standardized procedures.

Findings from the studies of English *et al* (English, 1995 #433; English, 1996 #434) showed that health workers could be trained to recognize standard agreed definitions of respiratory signs suggestive of severe disease. For the present study a mechanism was needed that would reliably identify signs and symptoms reflecting respiratory infection (cough, difficulty in breathing, rapid breathing for age) for the collection of nasal specimens and importantly to trigger referral of cases to the research clinic from the field for definitive evaluation. We therefore trained FWs for this purpose assisted by the same video recordings of patients used to train clinical staff. Assessment of the success of training in recording respiratory rates was indicated by an improved correlation between measurements by FW and by CO taken on the same child at the same time (where the CO result was the assumed "gold" standard). No specific inter-rater reliability analysis was carried out. Note that study FW were only charged with the responsibility of using the respiratory rate to make an assessment for referral to the clinic for a comprehensive review by CO of possible severe cases identified in the field (Appendix G3). Field workers referred a child seen at home to the clinic if they identified symptoms or signs reflecting a lower respiratory tract involvement, i.e. a history of acute cough or difficulty in breathing plus fast breathing (>50 per min for ages less than 12 months, and otherwise >40). The severity of respiratory disease was ascribed following review by a study CO, which would

include (a repeat) measurement of respiratory rate. A comparison of the two respiratory rate values showed significant correlation ($r=0.7$, $P<0.0001$). Further, an assessment of inter observer agreement on fast breathing using the kappa statistic was 0.371 representing fair agreement between the CO and FW. Of interest here is the sensitivity of the FW measurement of fast breathing assuming the clinician's assessment as the gold standard. Instances of misclassification (false negatives) occurred in 55% of the cases (Table 3.6). Admittedly this is quite high however this highlights the need for more than one method of referral in such studies - an issue that was taken into account in this study. At each home visit mothers were encouraged to attend the Research OP clinic if they observed any signs of respiratory infection at any time between FW visits, and asked to watch out for deterioration in the child's condition following a FW visit and requested to seek immediate help at the OP clinic in such an event.

Table 3.6. Correlation between Fieldworker and Clinicians' clinical classification

	Fast breathing FW		
Fast breathing CO	**0**	**1**	**Total**
0	7,161	647	7,808
1	757	613	1,370
Total	**7,918**	**1,260**	**9,178**

A proforma uniform to the research clinic and the paediatric ward was adopted to facilitate consistency between admitted cases and cohort cases. LRTI was assigned to children with acute cough or difficulty in breathing in association with any one or more of the following (i) raised respiratory rate for age, (ii) intercostal indrawing, or (iii) inability to feed, reduced conscious level, or hypoxia (O_2 saturation <90% by Oximetry), the latter group only if

confirmed by the clinician's own diagnosis of LRTI or bronchiolitis. Severe LRTI was assigned to a child with criteria (ii) or (iii), or both, of the above.

An analysis of digit preference of the respiratory rate measurement taken by the CO using a frequency distribution of the second digit did not indicate a preference for particular values (Figure). It should be noted that measurements were taken for half a minute and doubled to get the number of breaths per minute thus the second digit was always even. The frequency distribution plot shows no digit preference.

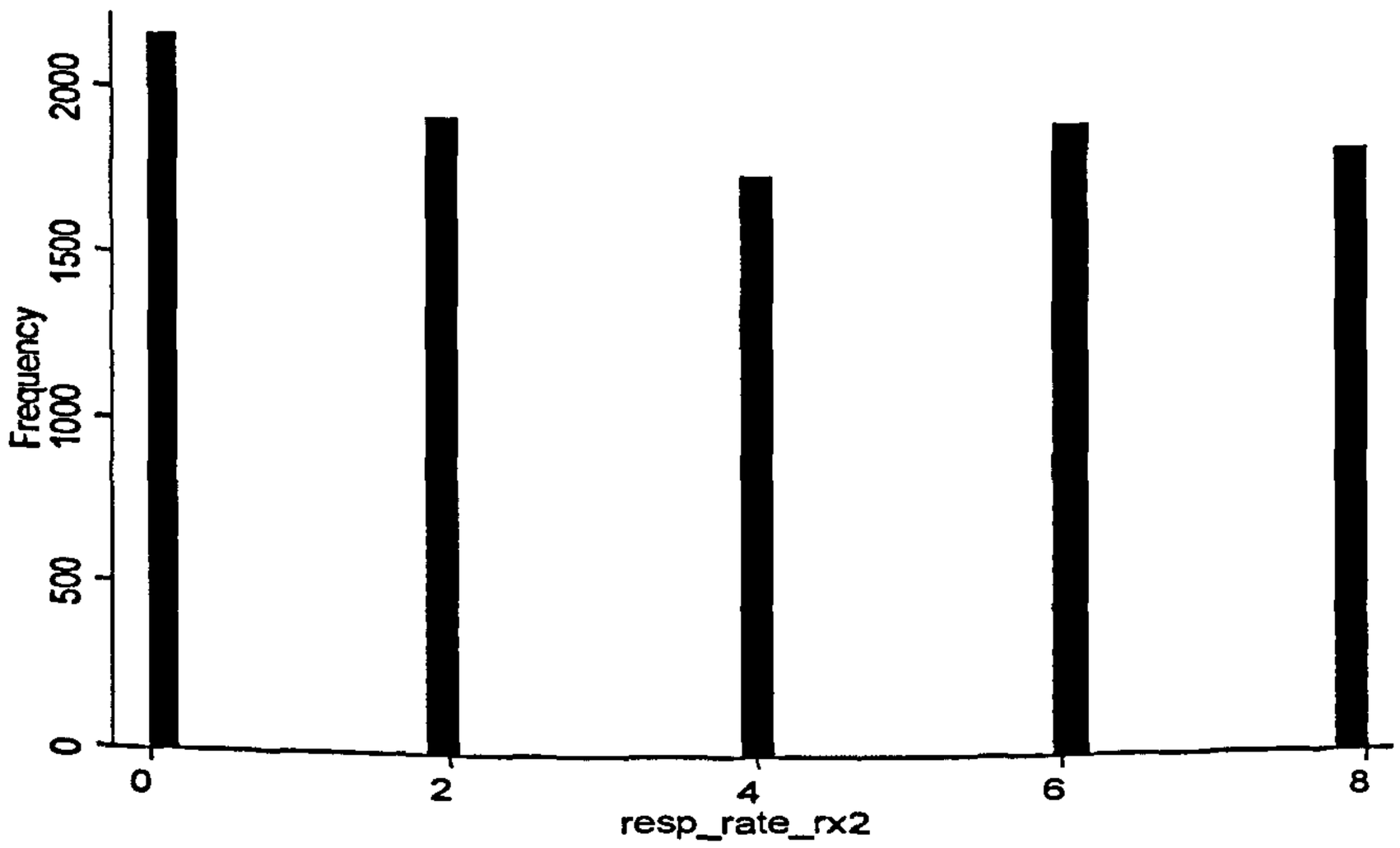


Figure 3.12 Frequency distribution of the second digits of respiratory rate variable

Lastly, as has been reported in numerous published studies (discussed in Chapter 2), RSV outbreaks are highly seasonal and usually appear in the rainy season in areas with tropical or subtropical climate [28, 39] such as that seen in Kenya. However, data from the present study suggests that epidemics are independent of rainfall, and were observed to occur in the hottest months with lowest rainfall. As discussed in Chapter 2, the issues surrounding this seasonality (i.e. how it persists and what triggers epidemics) remain unclear. Data collected over a longer period would be required to establish a clearer understanding of the

mechanisms at work. Surveillance is ongoing in the IP ward in Kilifi and it is hoped that such data will greatly enhance present knowledge.

Chapter Four

Evaluation of the dynamics of RSV-specific antibodies in oral-fluid samples

4.1. Introduction

Viral-specific antibodies provide information on the infection status of the host, and are particularly of use in support of clinical surveillance information (discussed in Chapter 2). The presence of anti-viral immunoglobulin M (IgM) is an indicator of recent infection. The presence of virus-specific immunoglobulin G (IgG) antibodies provides evidence of passive maternally-transferred antibodies or a past or present acquired infection. More specifically, paired samples from an individual which show a change from specific-IgG absence to presence (known as seroconversion) or a significant rise in viral-specific immunoglobulin G (IgG) antibodies (conventionally 4 fold) provides evidence of a current infection which could be a primary or a re-infection. Except where maternal antibodies persist to interfere with the immune response [88, 93], serial sampling for the determination of serum RSV-specific antibodies has yielded useful information on occurrence of infections [12, 19, 101] and is seen to be of particular use in age groups where clinical signs of infection may go unnoticed by surveillance due to decreasing severity of RSV infection with age or in those re-infected [12]. Furthermore, specific-antibody methods are of importance where the collection of nasal specimens is not easily undertaken, for example in older children and adults.

As previously described (Chapter 2), it is not always possible or desirable to collect blood samples by which to determine serological status, and oral-fluid sampling has shown promise as an alternative [20, 218]. It has been shown that oral fluid collected from around the gums is rich in exudated serum antibodies and thus reflects the antibody profile within serum [224, 225]. This has been shown for other viruses [212, 218] but not specifically for

RSV. The extent to which these data are analogous to RSV dynamics is unknown as a single study only has investigated anti-RSV antibodies in oral-fluid [20].

One of the key objectives of the household study was to determine and compare the incidence rates of RSV infection in different age groups. Birth cohort children were being serially bled to provide serological evidence of infections, but ethical approval for the same for older children had not been sought. Instead approval was sought and limited to the collection of oral-fluid samples at 3 monthly intervals from all family members, children and adults. This was to complement clinical sampling to identify possible missed asymptomatic infections in other household members thus overcoming the limitations of our clinical surveillance methods i.e. nasal washing collected only when symptoms were present, as well as to identify serological infection in adults. The oral-fluid was to be used to monitor the changing levels of IgG during the study with the incidence of RSV determined by examining seroconversion or boosting of anti-RSV antibodies in oral-fluid samples [20] from pairs of samples separated by a 3 month interval.

Existing data on oral-fluid antibody detection in relation to RSV and how this correlates with serum antibody ("gold standard") is, to the best of my knowledge, non-existent. The single study that has used oral-fluid to determine incidence of RSV infection in a population of children [20] did not make any comparison of antibodies in oral-fluid to serum. To overcome this problem in this study a comparison between serum and oral-fluid was included in the study design.

Presently, no known data detailing time-related changes or individual level variation in anti-RSV IgG antibody in oral-fluid exists. A subsidiary aspect of the three monthly oral

fluid collection surveillance design was therefore to characterize the temporal dynamics of RSV-specific antibody in oral fluid and its variability. The aim was to aid the interpretation of the 3-monthly interval data and subsequently to estimate re-infection rates. Correct interpretation of this data requires accurate measurement of each individual's exposure to RSV infection within a particular 3 month interval.

For the accurate interpretation of antibody measurements from sequential three-monthly oral-fluid samples, there are three aspects of antibody dynamics where knowledge is lacking (Figure 3.1). These are, i) the rate of natural background decay of IgG RSV specific antibody (r_1), ii) the threshold for variability in antibody measurement beyond which an observed antibody level can be assumed to be due to the boosting effect of re-infection within the sampling interval and iii) the possibility of an initial decay (r_2) immediately following antibody boosting (b) after infection at a rate faster than the background decay rate $r_2 > r_1$ (as has been seen in the case of measles [254]). In the absence of any information on the average level of boosting and subsequent initial rate of decay in RSV-specific antibody, it is difficult to ascertain what proportion of infections occurring at various time points in an observational interval (3 months) are likely to be followed by antibody decay to levels below the predefined criteria for assigning an individual as being re-infected.

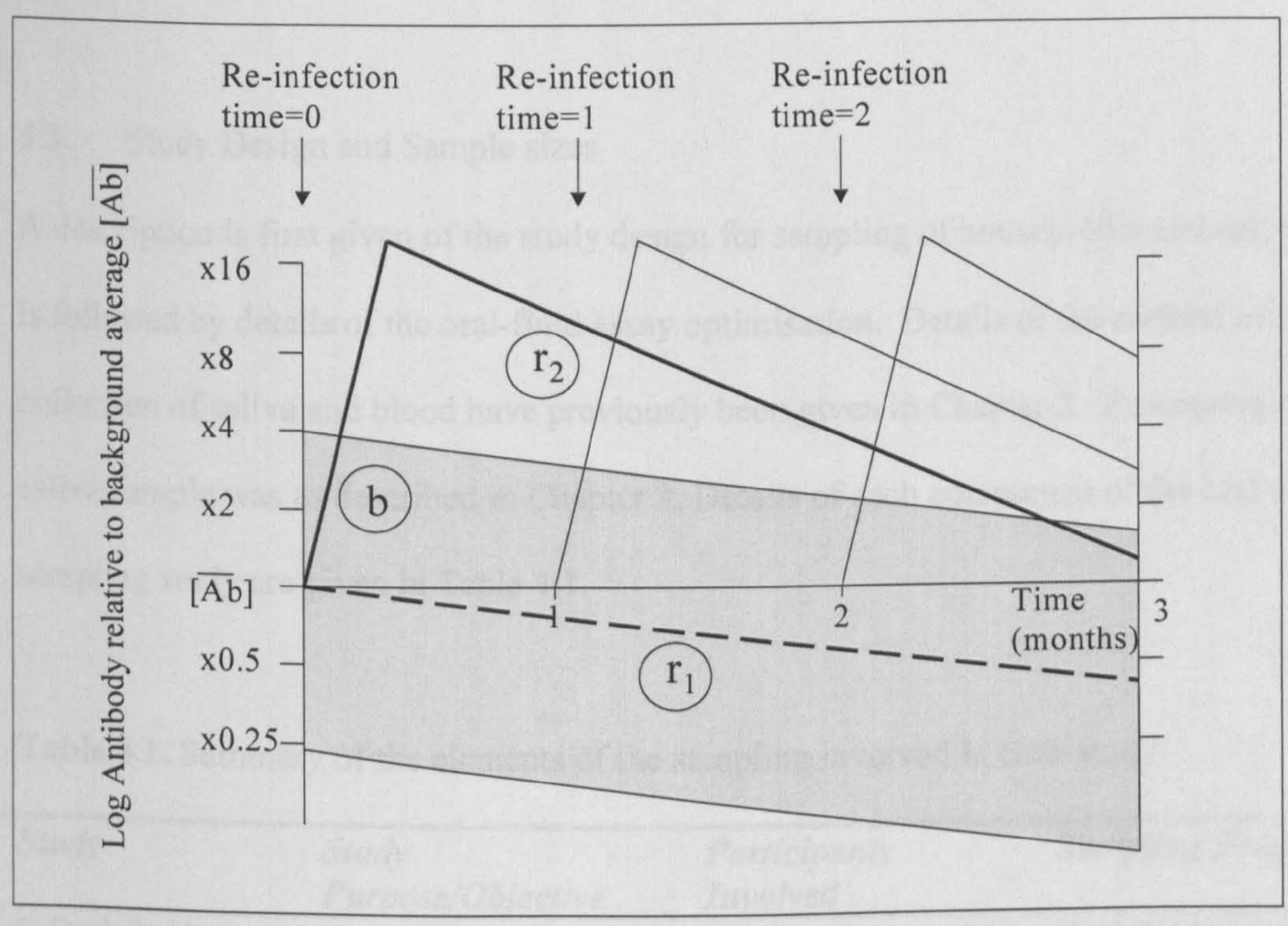


Figure 4.1. Schematic of antibody dynamics (log scale) of RSV infection over a 3 month interval, with or without re-infection. r_1, background rate of antibody decay in absence of boosting effect of re-infection. b, boosting effect following re-infection, and r_2 rate of decay following re-infection or boosting (b). Vertical axis represents average background antibody level [Ab] or multiples thereof. Grey zone indicates region of variability beyond which a measured antibody level is indicative of boosting (assuming 4 fold rise in antibody indicates infection.)

4.2 Chapter Aims

This chapter describes the study design of an oral-fluid surveillance system (introduced in Chapter 3) which complements clinical surveillance of RSV infections within the household. A supporting study is described which has the objective of improving our understanding of oral-fluid antibody kinetics, important to the interpretation of sequential 3 monthly data. An account is given of the enzyme-immunoassay used to determine RSV-specific antibody in oral fluid. The limitations of the ELISA method and results are discussed, together with the implications for the interpretation of household infection data.

4.3. Study Design and Sample sizes

A description is first given of the study design for sampling of household members, which is followed by details of the oral-fluid assay optimisation. Details of the method of collection of saliva and blood have previously been given in Chapter 3. Processing of the saliva sample was as described in Chapter 3. Details of each component of the oral-fluid sampling study are given in Table 4.1.

Table 4.1. Summary of the elements of the sampling involved in each study

Study	*Study Purpose/Objective*	*Participants Involved*	*Sampling Frequency*
i) Oral-fluid surveillance.	Complementary surveillance system. Used to estimate age-related re-infection rates and identify missed infection.	All household cohort participants, i.e. birth cohort, siblings and adults	Oral fluid samples taken at enrolment and every three months thereafter.
ii) Serum oral-fluid comparison study.	Assess performance of oral-fluid assay against assumed gold standard (serum).	Household members of the first 40 recruited households (convenience sample).	Paired blood and saliva samples taken twice over a 3 month interval. Samples collected from subgroups of 10 households in sequence.
iii) Antibody dynamics study.	Quantify RSV-specific IgG antibody dynamics.	30 individuals in each of three age groups; 0-5, 6-14 and 15+ randomly selected from within the household cohort.	45 individuals sampled weekly for 13 weeks within epidemic. Another 45 individuals sampled weekly in the inter-epidemic period for 13 weeks.

Three-monthly oral fluid surveillance

Oral-fluid samples were collected (see Chapter 2) from all members of each household in the main study at enrolment and thereafter samples were collected at intervals of three months until the end of the study. The objective of this study was to determine age-specific incidence of RSV re-infection within household occupants.

Serum Oral fluid comparison study

Paired blood and oral fluid samples from all (whenever possible) occupants of 40 households were collected on two consecutive occasions (at 3 month intervals), providing material by which to assess the performance of the oral-fluid assay for the detection of anti-RSV antibodies, in comparison with serum (assumed gold standard). 40 families were each sampled twice. For convenience, households selected for this study were the first 40 recruited into the household cohort. Two consecutive blood samples (at 3 month intervals) were subsequently collected from each of 4 subgroups of 10 households, sequentially, to compare antibody dynamics over the three month period at different time points. Blood samples were collected on the same date as the scheduled routine oral-fluid sample to facilitate comparison.

Antibody dynamics study

Intensive oral fluid sampling at weekly intervals was carried out for a period of 3 months (13 weeks) in a sub-sample of individuals within the participating households with the aim of quantifying RSV-specific IgG antibody dynamics over this critical time interval. Sampling was undertaken in 30 individuals in each of three age groups, 0-5, 6-14 and 15 and over years, stratified equally among epidemic and inter-epidemic periods.

Additionally (i.e. in addition to the 30 in each age class), any child from within the enrolled households identified as RSV antigen positive was enrolled into the study and sampled weekly for 3 months. It should be noted that the study was designed to include individuals experiencing primary and re-infections by which to evaluate differences in their antibody kinetics. Further details of the sampling procedures for this study follow.

i) Non-RSV cases

In September 2003 (inter-epidemic), 15 individuals within each of three age classes, 0-5, 6-14, and 15+ years were randomly selected from households using a block random design [255] to ensure equal numbers in each age class. Only one person per household was recruited. These individuals (total 45) were followed for 3 months (end December 2003) with samples of oral fluid collected at enrolment and every week thereafter (a maximum of 13 samples per person). In February 2004 (within epidemic season), a second sample of 15 individuals in each of three age classes, 0-5, 6-14, and 15+ years was selected and followed weekly for 3 months in a similar manner to that during the RSV season (i.e. expected number of samples 585). Sampling was necessary both within the RSV epidemic where serological identification of cases (by which to estimate boosting and early decay) would be highest, and between epidemics investigate any difference in the background decay rate (r_1 of Figure 3.1) relative to that within the epidemic. Based on previous family studies of RSV infection [16], it was anticipated that approximately 25% of household members would show serological evidence of exposure during one season, i.e. 11 cases (not taking account of cases identified by antigen positives – see below.)

ii) RSV cases

During the epidemic spanning 8/1/2004 – 2/6/2004 any child identified to be shedding RSV antigen (and not already in the intensive sampling study) was recruited into the intensive sampling study, as soon as possible after the laboratory result became available. An oral-fluid sample was collected at recruitment and each week for a further 13 weeks. The source of RSV cases can be separated into (a) birth cohort children, and (b) their siblings (aged<15 years) from enrolled households. Children already enrolled into the intensive shedding study and testing Ag positive were followed-up for a complete 13 weeks from this point. Assuming an average of 2.3 siblings to each cohort infant (estimated for households of the main birth cohort) it was projected that there would be a total of 160 siblings in the HH study. Roughly 20% incidence per annum was observed by antigen shedding in the infant cohort in 2002 and the assumption of half this rate (10%)[15] during the 2004 epidemic in the siblings would result in recruitment of about 16 infected individuals into the intensive oral-fluid surveillance. It followed that among the birth cohort infants (projected number 70) would give rise to 14 primary infections for intensive follow up. In total it was therefore expected that 41 infections would be identified from clinical surveillance and Ag testing (30) and serological monitoring (11, see previous section).

Results of sample collection

The results of samples collected from these three elements of serological surveillance are documented in Table 4.2. There were more antigen cases than was predicted 30 vs. 75. with relatively few instances of missed samples (< 15% in all stud arms). There were no instances of documented refusal to give repeated samples except in some (3) Muslim homesteads during the whole month of Ramadan for religious reasons - the belief that nothing (in this case Oracol) should be put in mouth.

Table 4.2. Total numbers of samples collected for each of the 3 elements of serological study.

Study		*Number of samples projected*	*Number of samples collected*	*Reasons for missed samples*
i) Oral-fluid surveillance.		4023*	3536 (88%)[$]	i) traveling ii) loss to follow-up
ii) Serum oral-fluid comparison study.		434	374 (86%)[$]	iii) refusal to give oral-fluid sample during the Muslim fasting holiday (Ramadan)
iii) Antibody dynamics study.	In-season RSV Ag positive cases	585 390[&]	493 (84%)[$] 975 (75 positives)	
	Off-season	585	520 (89%)[$]	

* from approximately 24 months of follow-up (the median follow-up time). (9*447)
[&] 30 Ag positive cases expected
[$] Proportion of expected number of samples

4.4. Oral-fluid Enzyme Linked Immunosorbent Assay (ELISA) and Optimisation

Oral-fluid specimens were analyzed for anti-RSV antibody using a published method [20] to quantify RSV specific IgG. The assay comprises two stages. First, the concentration of total IgG is determined using an IgG capture assay. For antibody capture assays very low concentrations of IgG are required to saturate the solid phase. The process can therefore be successfully used with saliva. A set of standards of known concentrations are included in this step from which the concentration of the sample is calculated. Second, following from step one, a quantity of oral-fluid sample calculated to contain a standard concentration of IgG (1mg/L) is used in an indirect sandwich anti-RSV ELISA. These two stages are described further in the following sections and the detailed laboratory methods are provided in Appendix M (p 342).

Step 1: Antibody capture assay (Fig.4.2)

A solid support (microtitre plate) is coated with anti-IgG antibody. When the test specimen is added, IgG class antibody contained within is captured giving (after washing) a sample free from other components. The presence of the antibody-antibody complexes is revealed by the addition of an enzyme-conjugated antibody specific to the captured immunoglobulin class, which catalyses a suitable substrate added to produce a colorimetric product.

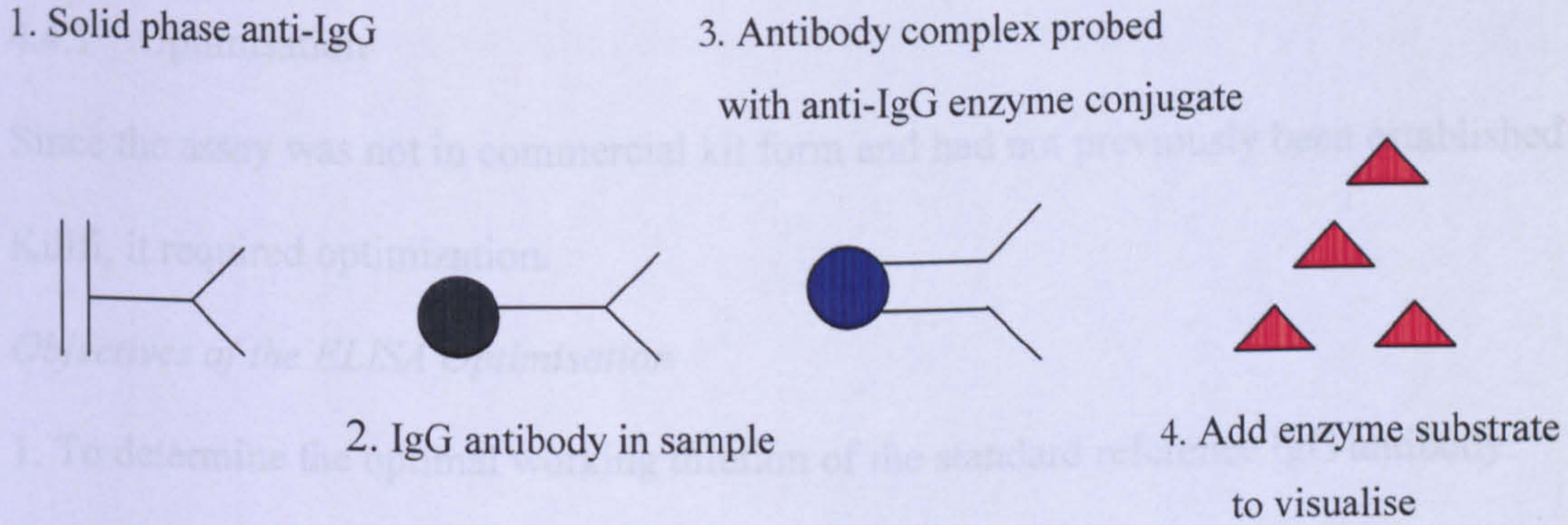


Figure 4.2. Illustration of stages (1-4) of the IgG Capture ELISA assay [256]

Step 2: Anti-RSV assay (Fig.4.3)

The oral-fluid specimen was diluted to a standard concentration of total IgG, following Step 1, to overcome variation in total IgG associated with crevicular fluid sampling. Virus specific antigens (whole cell extract) are immobilised on a solid phase to which the oral-fluid specimen is added. Any antibody in the sample against the viral antigen is bound. This binding is revealed by the addition of an antibody against the IgG class of interest conjugated to an enzyme that generates colour upon the addition of a substrate.

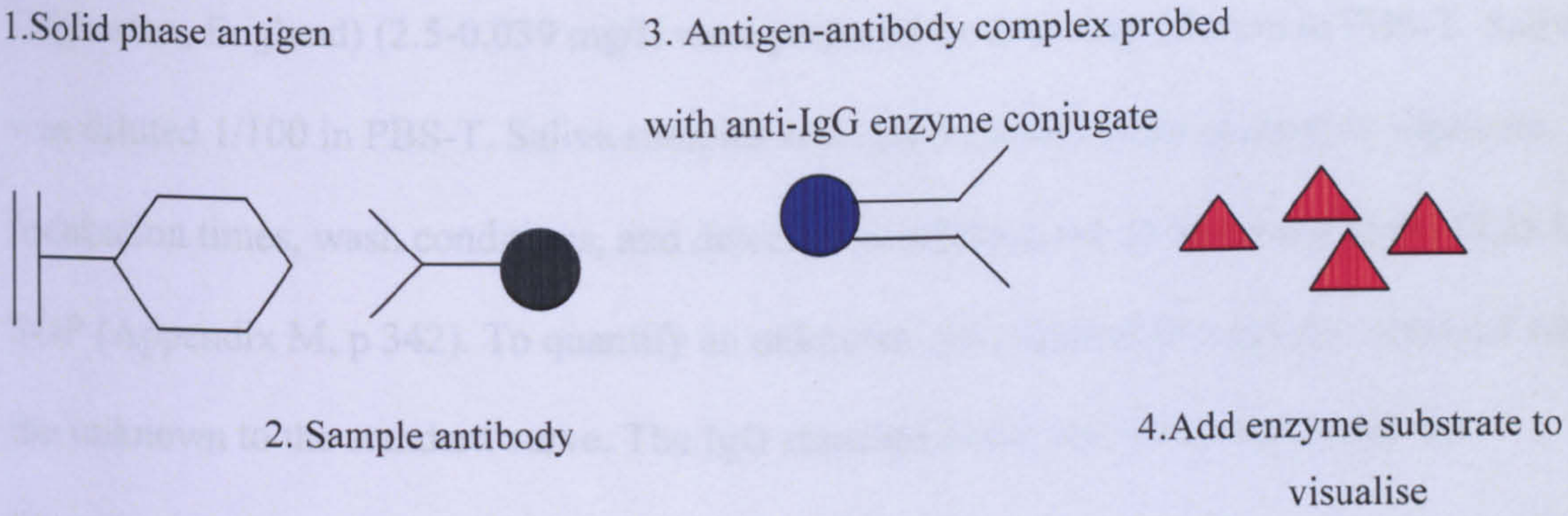


Figure 4.3. Illustration of stages (1-4) of the direct ELISA assay[256]

4.4.1 Optimisation

Since the assay was not in commercial kit form and had not previously been established in Kilifi, it required optimization.

Objectives of the ELISA Optimisation

1. To determine the optimal working dilution of the standard reference IgG antibody.

2. To determine the optimal working dilution of the test sample.

3. To determine the optimal coating concentration of RSV antigen.

4. To investigate the sensitivity of the assay at low and high titres of anti-RSV antibody.

5. To evaluate the method against assumed gold standard serum since the only previous publication did not compare oral fluid with serum.

a) Determination of the optimal working dilution of the standard reference IgG antibody for the IgG capture assay

A series of experiments were performed to determine the linear limits of the standard reference curve. Flat-bottomed 96-well plates were coated overnight at 37°C with 100 ml/well rabbit antihuman IgG diluted 1/3,000 in sodium carbonate coating buffer. The plates were washed four times in phosphate-buffered saline (PBS) containing 0.05% Tween 20 (PBS-T). The plates were then blocked with 5% dried milk powder in PBS-T (200 ml/well) at 37°C for 60 minutes. IgG calibrant standards (180μg/ml IgG, Binding Site,

Edgbaston, England) (2.5-0.039 mg/l) were prepared by doubling dilution in PBS-T. Saliva was diluted 1/100 in PBS-T. Saliva samples and IgG standards were assayed in triplicate. Incubation times, wash conditions, and detection conditions are as described in the ELISA SOP (Appendix M, p 342). To quantify an unknown, we compare the activity obtained with the unknown to the standard curve. The IgG standard curve was plotted to check for linearity.

Result

There was very good reproducibility of the replicates and log-linearity through out the standard dilution range used.

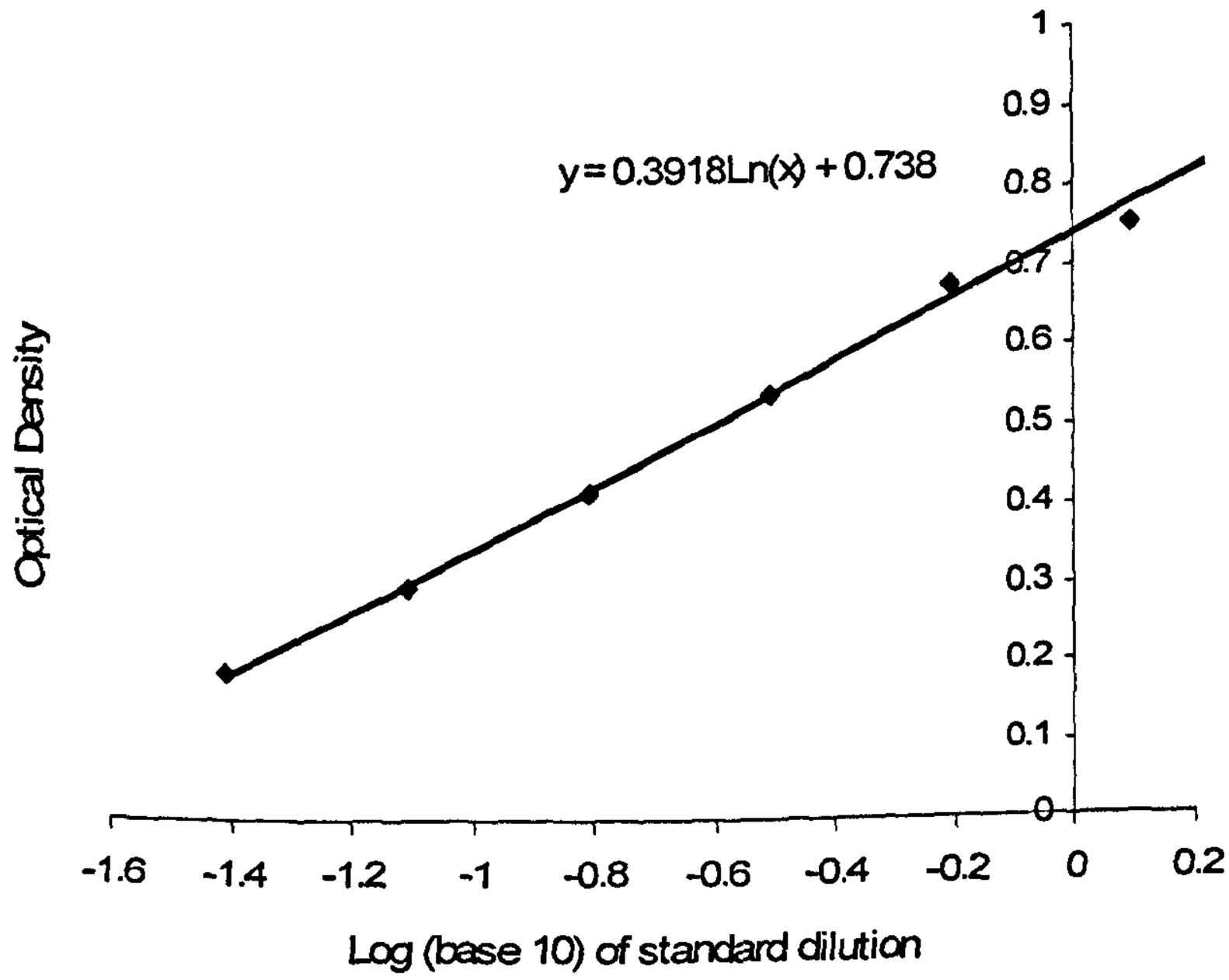


Figure 4.4. IgG standard curve

The IgG concentration of an unknown, x, was derived from standard curve as follows. We assume a linear relationship between the logarithm (base 10) of x and optical density, OD,

$$OD = a + b(\log x) \qquad (4.1)$$

where a and b are the intercept and slope of the regression line, respectively. Rearranging 4.1 we obtain the following expression for x,

$x = 10^{(OD-a)/b}$.

b) Determination of the optimal working dilution of the test sample

We investigated the optimal working dilution of samples. There was no documented reason for using the 1/100 dilution adopted in the reference method paper [20]. Twelve 3-monthly oral-fluid samples from children in the first, second or third year of life, and some convalescent samples were assayed in triplicate, to improve accuracy, using doubling dilutions from 1/10 to 1/20480. To confidently choose a working dilution samples were taken from individuals of different ages; children of different ages assumed to have differing infection history which would be reflected in the levels of total IgG. One sample each from 3 children in each of three age class- 0, 1, 2 years and 3 children from the IP study (< 5 years) (Refer to section 2.12) were tested and the saturation zones and linear portions of the dose-response curves defined. Ideally the estimated concentrations from absorbance values falling within the linear limits of the curve should be consistent after taking account of the dilution factor.

Results

The linear region of the dose response curves fell within the 1/20 and 1/640 dilution range (Figure 4.5) hence; either the 1/100 or 1/200 dilution would be good working dilutions. The results show some variability in amount of total immunoglobulin depending on which dilution was used and it was not apparent which of the two dilutions would be better. Two options are available. i) Use both dilutions to assay samples and take an average of the two. However, this would increase the work load. ii) Choose a single dilution but dilute the sample further if absorbance readings were outside the linear limits of the standard curve.

The second option was adopted i.e. use the 1/100 dilution allowing for further sample dilutions if needed. This resulted in minimal increased work load.

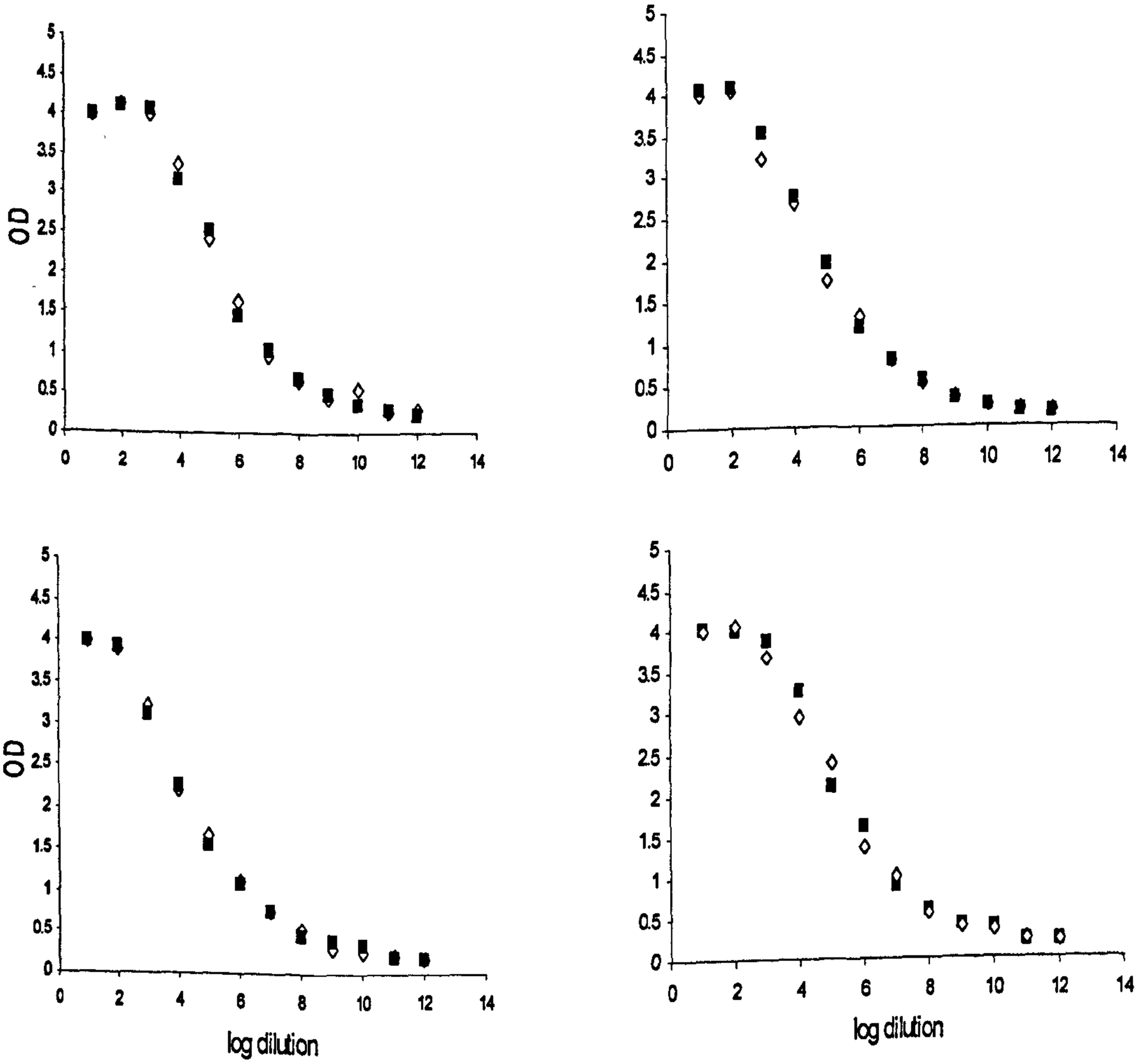


Figure 4.5. Absorbance readings for total IgG using doubling dilutions from 1/10 to 1/20480 (identified as labels 1-12) for different oral-fluid samples: 2 birth cohort child, 9 and 12 months (right) and acute and convalescent samples from IP children, 3 and 3.3 years (left). OD is optical density at 492nm. Includes two replicates of each sample (open diamond and filled square markers).

c) Optimisation of the RSV specific assay for oral fluid

The following series of experiments were conducted to optimise the assay for detection and quantification of RSV-specific IgG. Briefly for each experiment, six columns of a ninety-six-well plate were coated with RSV-A2-infected lysate; the other six coated with mock-

infected cell lysates as describe in the SOP (See Appendix N, p 344). Infected and uninfected lysate was grown at Birmingham University Medical School by Dr Paul Scott. Plates were coated in Birmingham and shipped on ice to Kilifi.

1.) Initially, 6 oral-fluid samples from adults in Kilifi were used because they should have high level of specific IgG from previous infection. A pooled serum sample (mixed serum sample form several adults in Kilifi) was included as control. Samples were assayed in duplicate from 1/10 to 1/20480.

Results

Low absorbance readings (OD<0.2) were observed with negligible difference between positive signal (response to infected lysate) and noise (response to mock antigen). Further experiments were carried out using 5 samples from cohort children with similar results.

2.) The logical next step was to change the sample dilution range to 1/2 to 1/64 to increase the specific antibody concentrations. Oral-fluid samples from two groups were assayed, i) 27 samples from individuals with known clinical infections i.e. acute convalescent sample pairs from in-patients (IP) expected to show antibody boosting or seroconversion (1-5 years), and ii) consecutive three month samples from the same child over the first year of life (samples from 3 children).

Results

Some acute convalescent paired oral-fluid samples (an example is illustrated in Fig. 4.6) showed a rise in antibody. However, most samples tested showed no detectable RSV specific antibody.

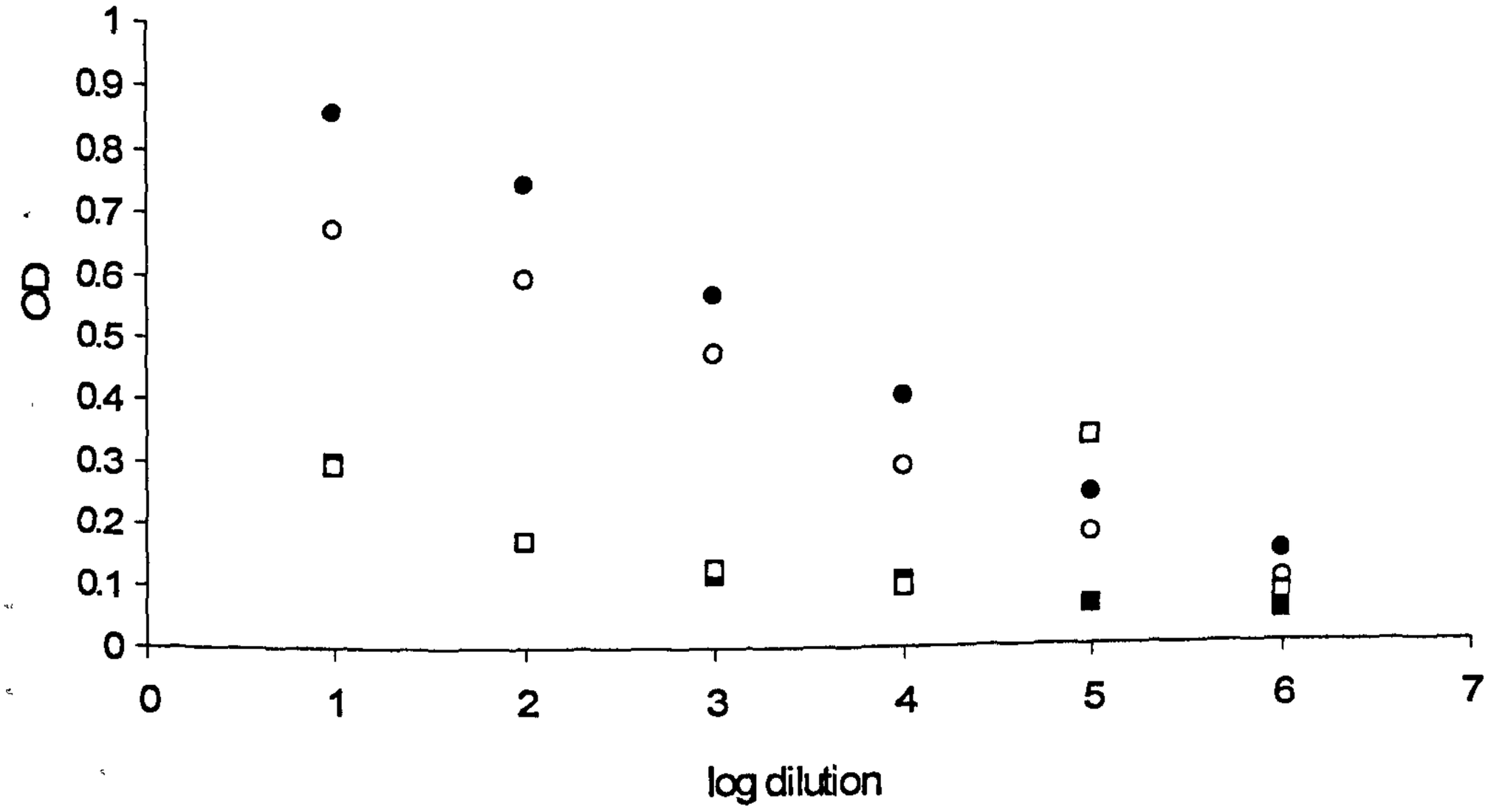


Figure 4.6. Absorbance readings from paired acute and convalescent oral fluid samples from Patient 56134, diluted from 1/2 to 1/64. OD is optical density at 492nm. Includes two replicates of the acute sample (open and filled square markers) and two replicates of the convalescent (open and filled circles).

There were two possibilities to explain the apparent variability in results; some samples were showing boosting of antibody between the acute and convalescent sample while others were giving no signal. The results could indicate that the assay worked and possibly that the samples with no signal did not have detectable levels of specific IgG. Alternatively, there was a fault with the assay such as insufficient blocking, inadequate sample volume used and/or nature of storing samples (issues explored below). To resolve this, paired oral-fluid samples of serum samples that had previously shown rising anti-RSV titre (boosting) were assayed (Serum sample results courtesy of Rachel Opiyo). These ideally would show the same antibody boosting observed in serum.

3.) The assay was therefore repeated using paired serum and oral-fluid samples from children into the second year of life. Full range of samples from nine children were assayed some results of which are shown in Figure 4.7

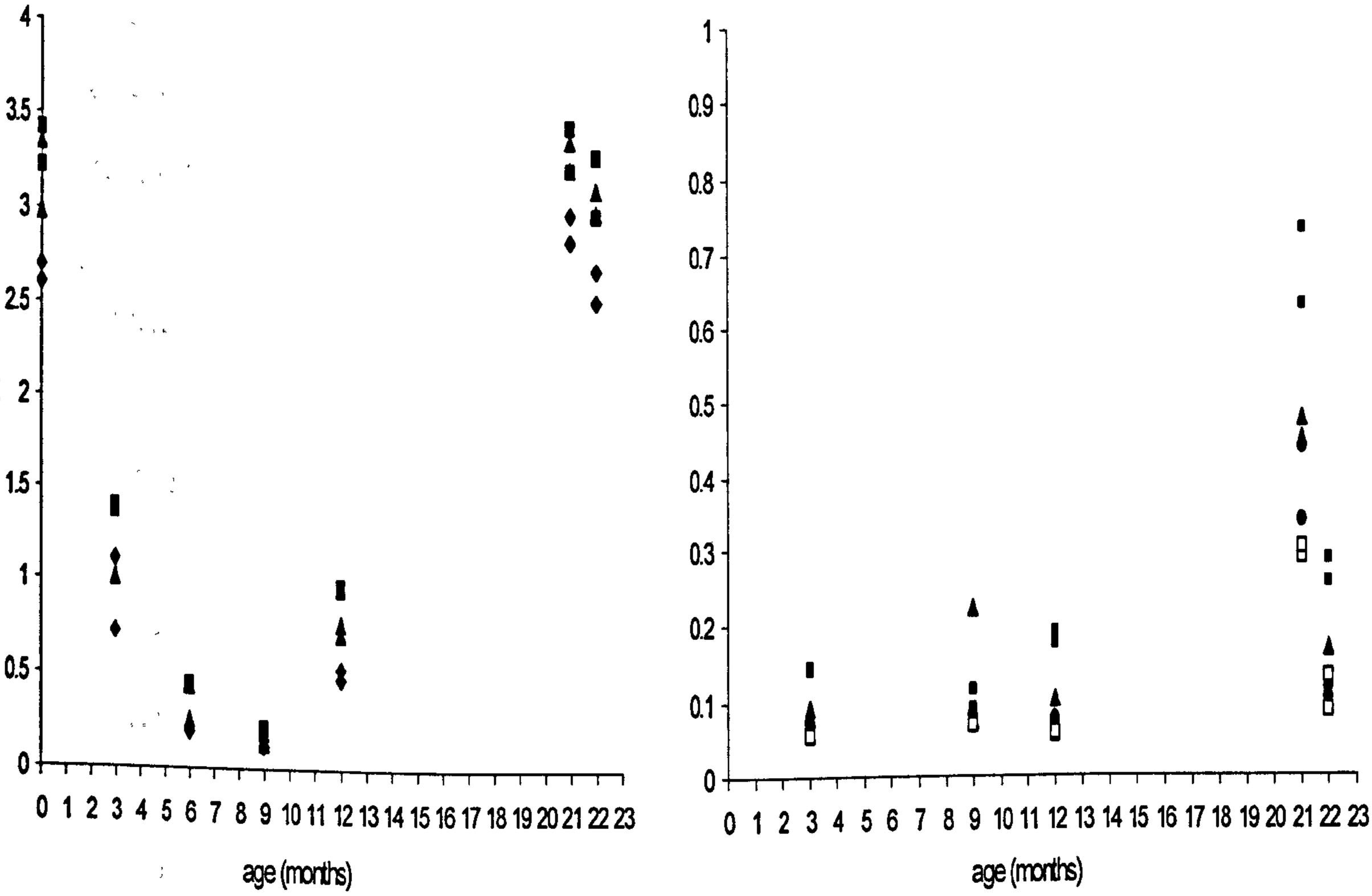


Figure 4.7. Age plots for serum (left) at 1/100 to 1/400 dilution and oral-fluid (right) at 1/2 to 1/16 dilution. Note the difference in scales used. OD is optical density at 492nm. The filled black boxes are 1/100 and 1/2 dilutions for serum and oral-fluid respectively, orange is 1/200 and 1/4, the blue 1/400 and 1/8 while the open boxes are 1/16 dilution of oral-fluid. Last two samples in each panel are acute and convalescent samples. Oral-fluid sample was not routinely taken at birth and the 6m sample was not collected.

A roughly similar antibody pattern as that seen in serum was observed in oral-fluid (Figure 4.7), but still at very low levels (negligible) especially after correcting for background (results not shown). The serum plot at 1/100 or 1/200 gives a good profile of decay (presumed maternal antibody) and subsequent rise in antibody concentration (following infection). The pattern of decay in maternal antibody is not observed in oral fluid. However, a rise in the acute/convalescent sera does appear. The assay was repeated using adult samples (presumed to have higher antibody levels in general) taken from the household study also giving very low absorbance readings (results not shown). Subsequent

repeat experiments were carried out using standardized saliva samples corresponding to 1mg/L IgG for the specific assay to correct for the inherent variability in oral fluid samples.

4.) To further investigate the observed results, various other troubleshooting options were explored in an attempt to improve the strength of the signal observed and reduce the background. For each experiment in this section samples from 8 adults were assayed.

i) Substrate processing time

In varying the substrate processing time from 10 minutes to 30 minutes it was found that the longer the substrate was left to process the higher the absorbance readings for both the signal and mock (noise) (Figure 4.8).

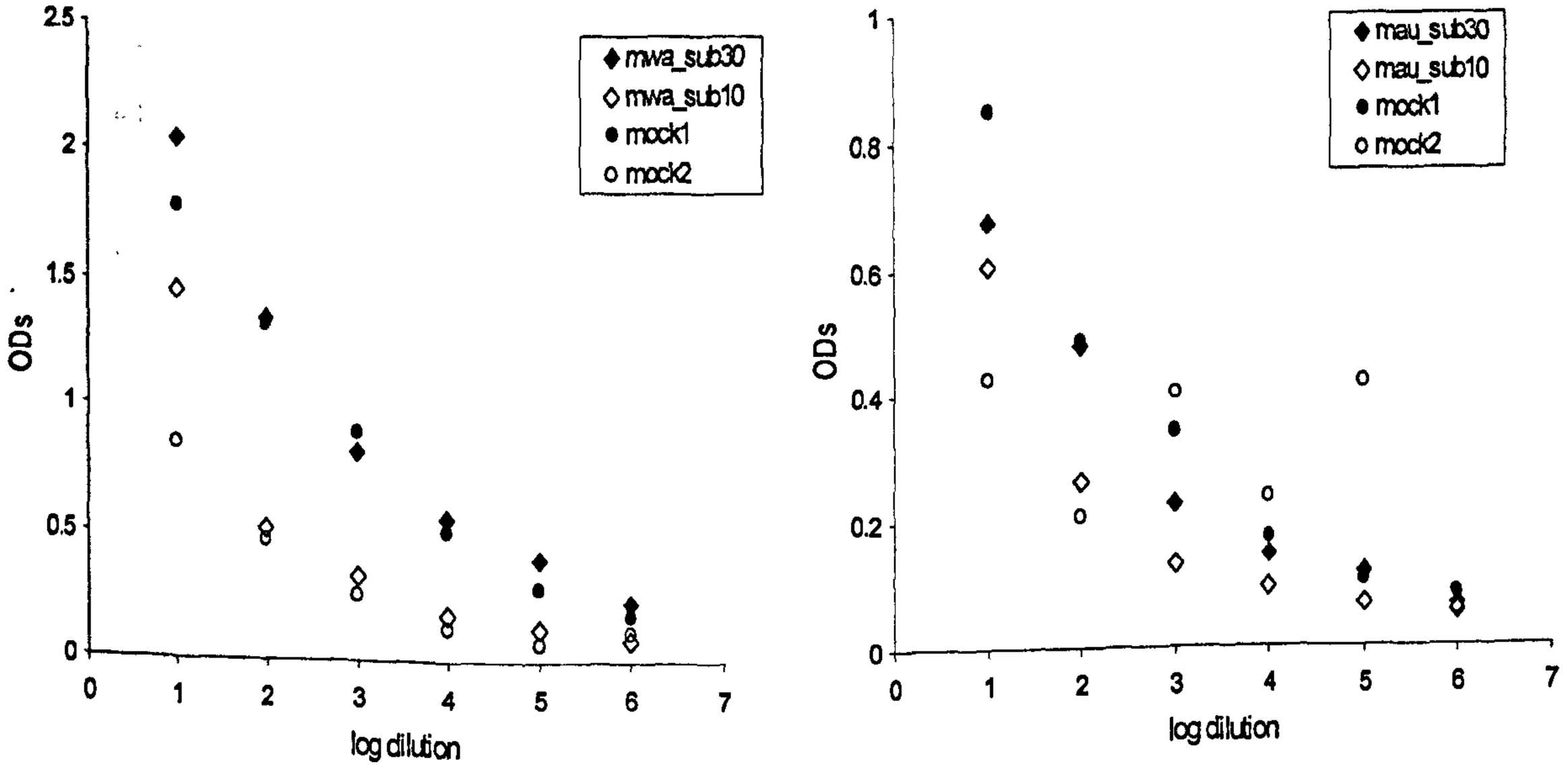


Figure 4.8. Comparison of substrate processing times 10 vs. 30 minutes using samples from two adults (Mwa (a) and Mau (b)). OD is optical density at 490nm

ii) Comparing fresh and frozen samples

Fresh saliva samples from various adults in Kilifi were used on the working assumption that the method of sample storage may affect the outcome observed. Sodium azide used in

the preparation of the freezing mixture has been known to inhibit peroxidase reactions. No appreciable differences in absorbance readings were observed (results not shown).

iii) Blocking Optimization

In an attempt to reduce the signal to noise ratio the use of a higher concentration of blocking solution was explored. Working on the assumption that increasing the block would reduce the background problem, plates were blocked using 5% (as per SOP) and 10% (new) block solution. The results from this experiment are shown in Figure 4.9.

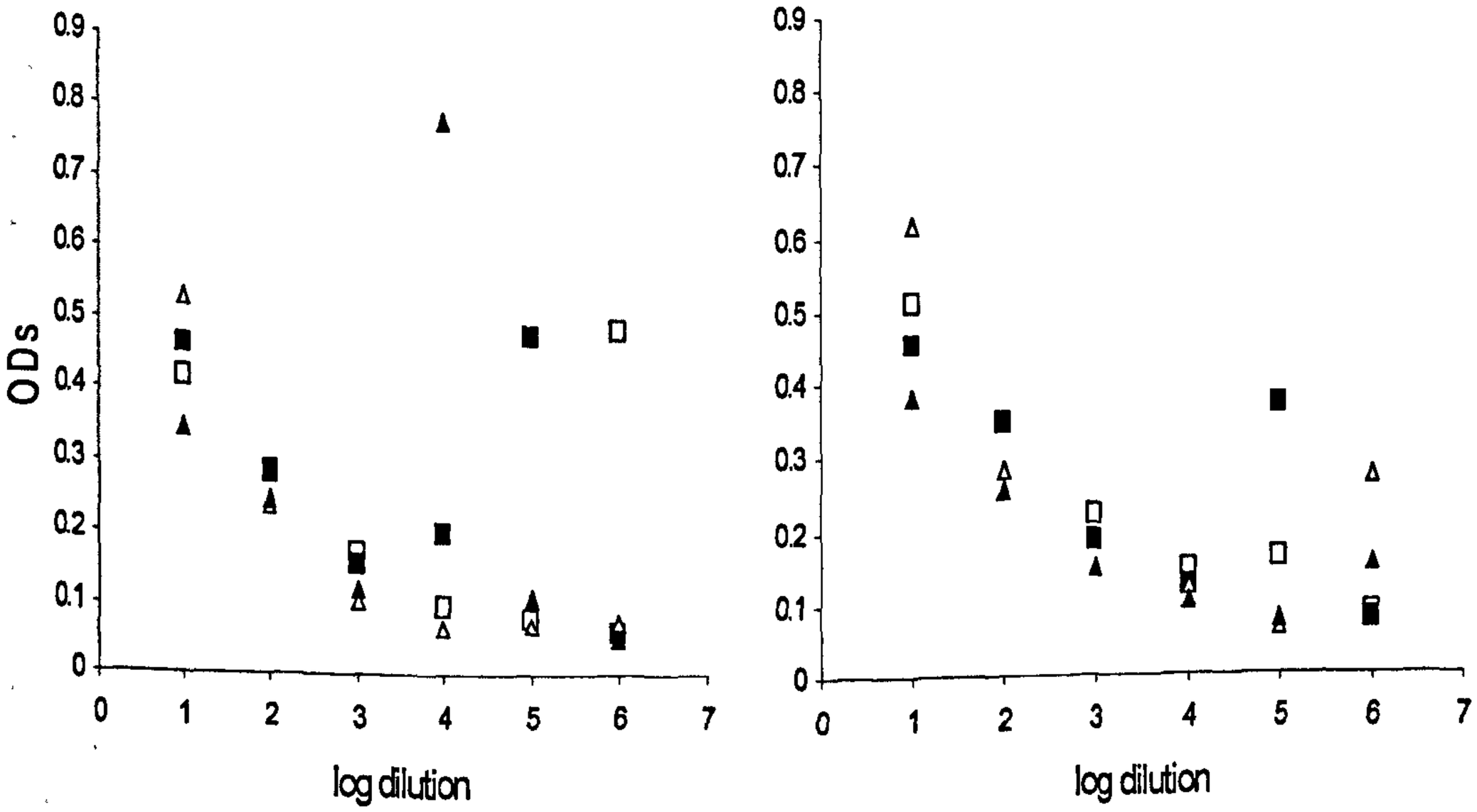


Figure 4.9 Comparison of absorbance readings using different concentrations 5% (left) vs. 10% (right) of blocking solution. OD is optical density at 492nm. Includes two replicates of the sample (open and filled square markers) and two replicates of the mock (open and filled triangles) for each assay.

iv) Comparing effect of different secondary antibody

Investigated the use of a different secondary antibody in the specific assay as a possible reason for the low absorbance. Wilson *et al* [20] used Horseradish Peroxidase (HRP)-conjugated goat anti-human IgG as the secondary antibody for the RSV specific ELISA

with no documented reason for doing this. In this study HRP-conjugated **rabbit** anti-human IgG was used in both the capture and the specific assay. An experiment was performed to investigate the difference between the two secondary antibodies. Oral fluid samples taken from Kilifi adults were assayed and serum included as the standard.

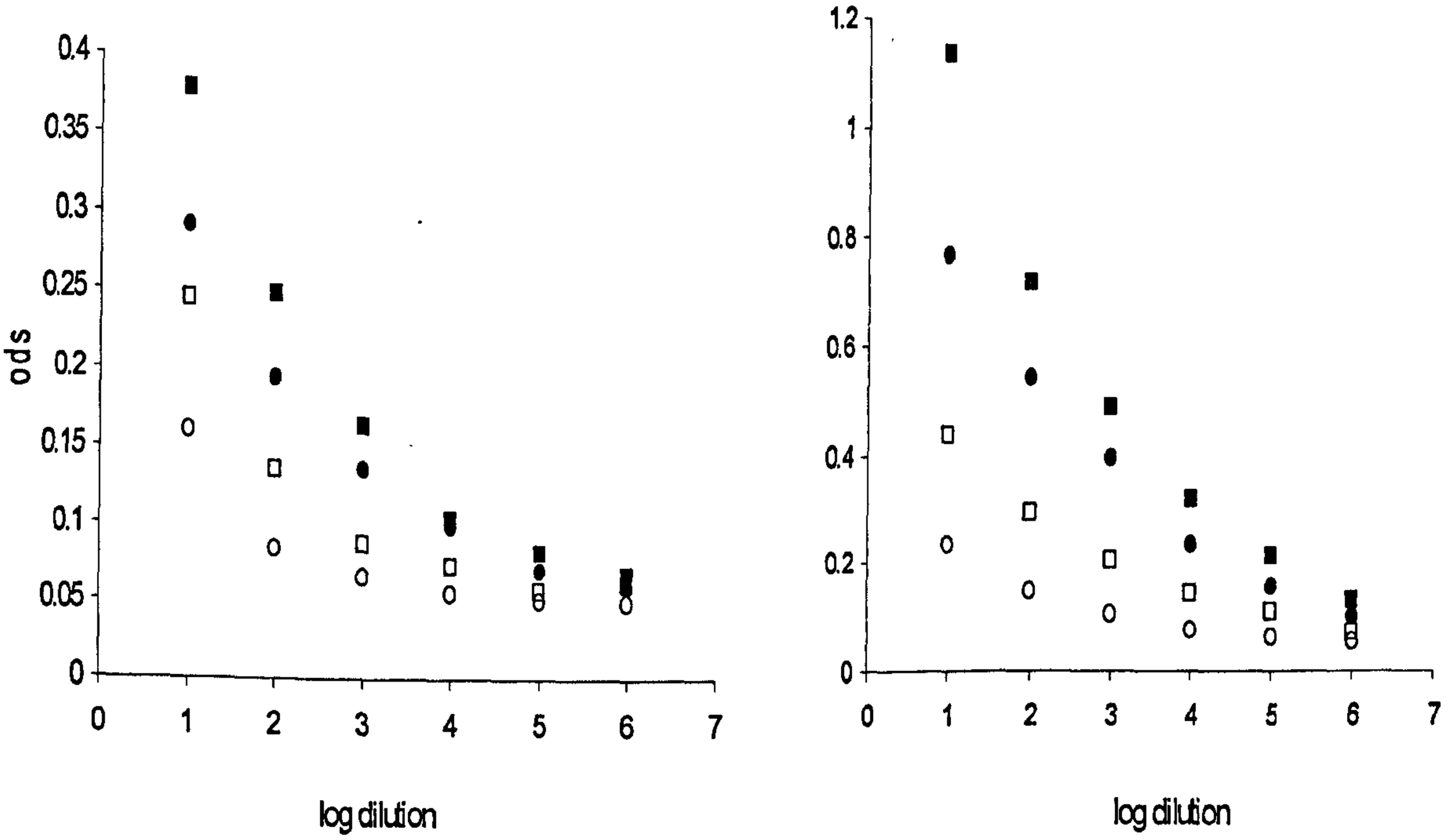


Figure 4.10. Comparison of absorbance readings using two different secondary antibodies on both serum (right) and oral-fluid (left) samples. OD is optical density at 492nm. Dilution used was 1/2 - 1/64 for oral-fluid and 1/50- 1/1600 for serum (shown on the graph as 1-6). HRP-conjugated goat antihuman (filled (signal) and open box (noise) markers). HRP-conjugated rabbit anti-human (filled (signal) and open circles (noise) markers).

From the example in Figure 4.10 above and in all samples analysed the HRP-conjugated goat anti-human IgG produced higher absorbance readings than the HRP-conjugated rabbit anti-human IgG [as used by Wilson]. Despite the fact that there is a difference in the results given by the different secondary antibodies, the HRP-conjugated goat anti-human also gave higher noise signal which seems to even out the overall effect. Still the HRP-conjugated goat anti-human was adopted for use. Subsequently, previously assayed paired serum and

oral-fluid samples were also tested using the new secondary antibody. It is now know that there was a problem with the antigen coated plates which has only recently been resolved by growing our own Ag and changing various aspect of the coating method (see discussion later).

Results

Similar dose response patterns were observed (some examples illustrated in Figure 4.11 and 4.12). In most instances the absorbance readings of saliva were found to be roughly half the magnitude of that seen in serum. Note that the sample dilutions used for serum and oral-fluid were different. Signal to noise ratio for oral-fluid sample is in several instances almost 1 i.e. no difference between antigen and mock.

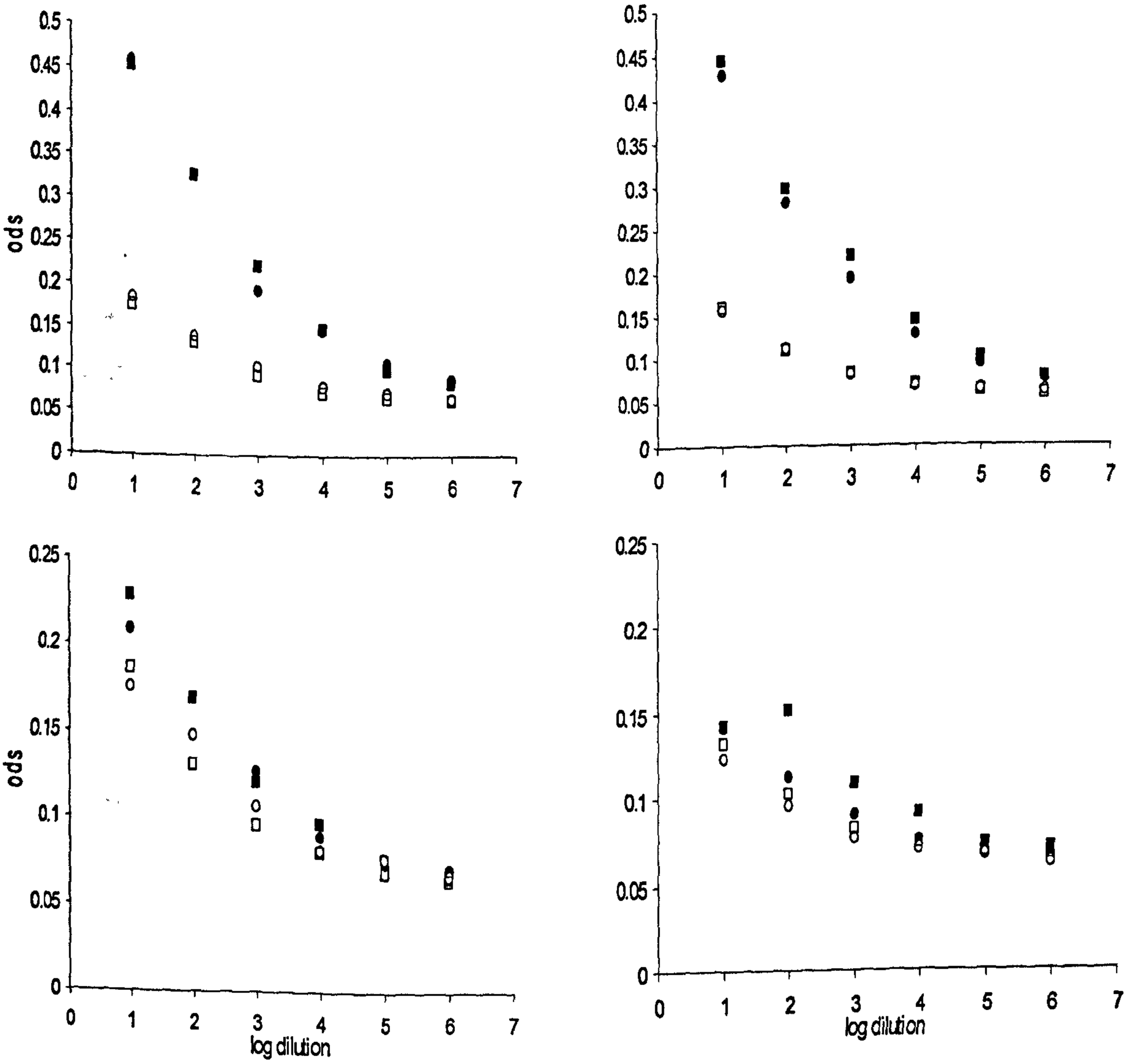


Figure 4.11. Absorbance readings for paired serum (top) and oral-fluid (bottom) samples from a household child. First set of samples on the left taken on 20-05-03 (6.9 years) and second sample on the right taken on 10-09-03 (7.2 years). Dilution used was 1/50- 1/1600 for serum and 1/2 – 1/64 for oral fluid. OD is optical density at 492nm. Samples run in duplicate (signal -filled box and circle markers and noise- open box and circle markers).

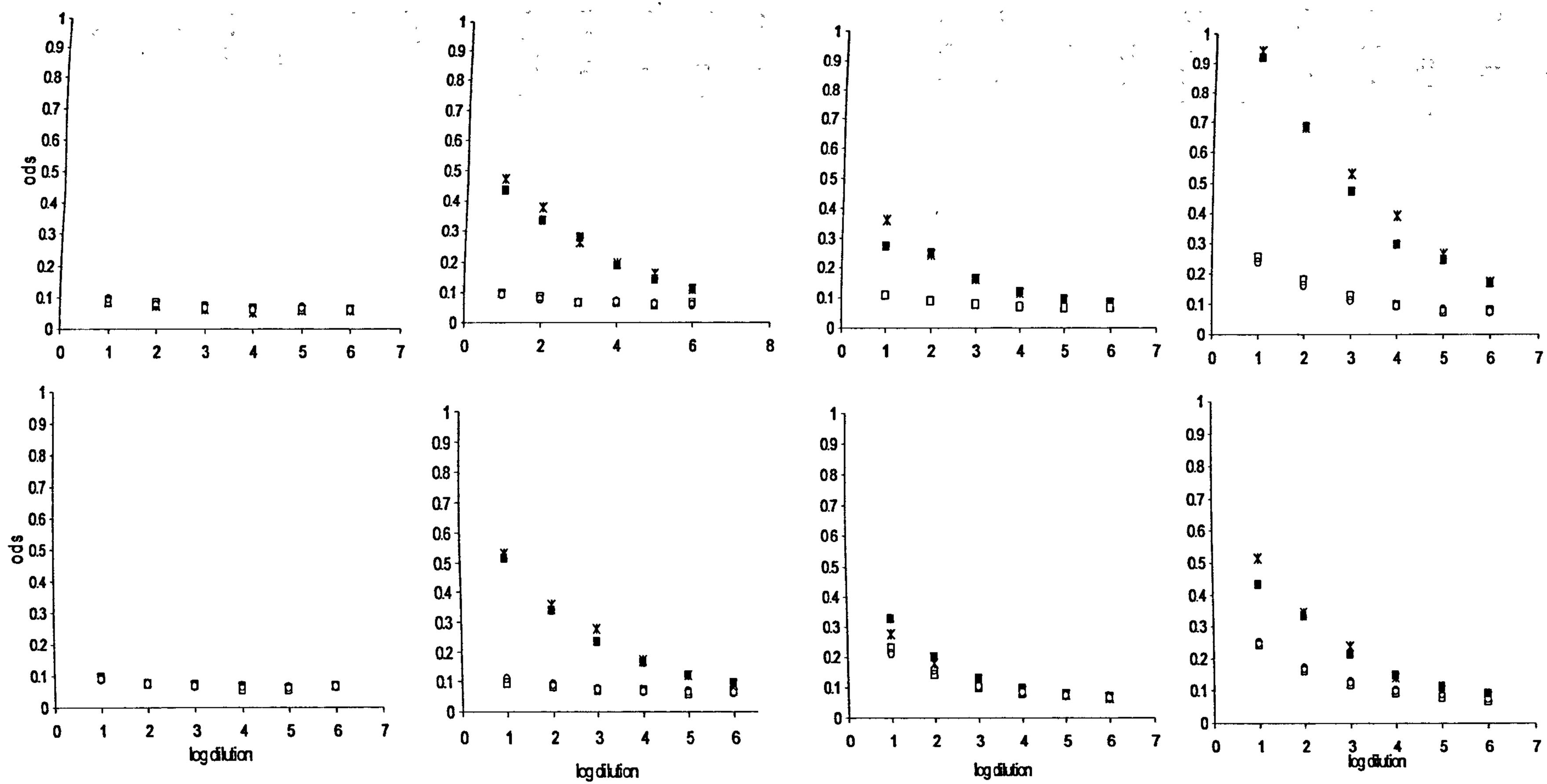


Figure 4.12. Oral-fluid (bottom panel) and serum (top panel) samples from patient taken on four consecutive dates (from left to right); 29-03-04, 27-04-04, 12-08-04 and 06-01-05 respectively. OD is optical density at 492nm. Two replicates (star and filled box markers) assayed for each sample each with mock (open circle and box markers). Dilutions as in Fig 4.11.

d) Determination of the optimal coating concentration of RSV antigen

With less than promising results, a review of the dilution of the antigen used was undertaken. Until this point, plates coated with 1/64 antigen dilution, the dilution that gave the best signal to noise ratio for serum had been used. Oral-fluid samples were titrated against antigen in a checker board analysis. The checker board analysis was done using the same previously used lysate grown at the Birmingham laboratory and new locally grown lysate. Results from this experiment show that the 1/64 antigen dilution used gives general low absorbance reading. The signal to noise ratio in most instances was ~1 except at the neat and 1/2 antigen dilution. This may explain why we were getting such low absorbance readings in general. Using the lysates undiluted or at a 1/2 dilution was not feasible in terms of lysate production.

4.4 Discussion

The serological assays have the added potential for the identification of possible asymptomatic cases missed by routine clinical surveillance. This is especially important in older children and adults with typically more mild illness of short duration or who in some instances show no symptoms (Chapter 2). Previous work in children indicates 50% sensitivity for RSV infections through clinical surveillance [19, 94, 144]; even less in adults [125]. Though less informative than viral detection (Chapter, Section 3.4), oral-fluid ELISA determined infections identified by changes in antibody titers would have picked up infections missed by clinical surveillance for a number of reasons, including asymptomatic infection, IFAT sensitivity particularly in adults and older children [125, 241]. However, we were unable to proceed with the oral-fluid assay due to unreliable signal the reasons for which are not clear.

Lower concentration of immunoglobulin in oral fluid makes the quality of the sample very important. The most important source of the IgG antibody in oral fluid is the crevicular fluid as previously indicated. It is therefore vital that the device used to collect oral fluid targets this fluid. There are various collection methods. A study [213] comparing several devices rated the Oracol sponge device the most comfortable and with significantly higher geometric mean titres of total IgG than for the other two devices (41.7 mg/L (95%CI 28.8-58.9) sponge versus 25.1 mg/L (18.2-35.5) OmniSal and 22.9 mg/L (18.2-28.8) OraSure). Assuming a direct relationship between virus-specific and total IgG, it is thus unlikely that the low signal in our experiments was as a result of the quality of the sample and the collection device used.

A similar procedure was, at the time of writing, giving very low signal to noise ratio with sera from cohort children (by a PhD student colleague, Rachel Opiyo). Results from the serum assay will be essential in understanding the problem(s) with the oral-fluid assay. After consultations with other researchers who had previous experience with the use of OF samples a joint decision was taken between my supervisors and myself in consultation with my appointed Student Advisory Committee to terminate this line of work and to concentrate of other aspects of the project. At the time, the major reason underlying this decision was the poor results obtained with the serum assay (considered the "gold" standard). It was felt that given these results it was highly unlikely that better results could be attained with the OF assay within the time constraints for completion of the thesis. It is thought that several aspects of the previous protocol were flawed contributing to the poor results. The protocol was based on a previously published method. A new improved assay protocol has since been developed (February 2007) for the serum assay details of which

and an interpretation of the possible problems as well as prospects for future work using the OF samples collected are discussed below.

Following several modifications to the experimental protocol the test signal-noise ratio for serum assay has improved significantly suggesting a similar result would be expected when used with oral-fluid samples. The modifications implemented include:

i) the viral antigen recovery method was modified. Previously, recovery of virus from culture was by the addition of detergent and shaking to rupture the cells. The detergent acts by interfering with the cell membrane and therefore aids in the breaking up of the cell membrane. It is thought that this method did not efficiently break up epithelial cells. This method was found to limit the amount of virus antigen recovered. In addition, the added detergent was found to prevent binding of virus antigen to plates. The method of viral recovery was therefore changed and cell break up is now done by sonication. Various combinations were tested before one was selected: sonication alone, detergent plus sonication and using detergent only. No difference was observed between the two methods using sonication but there was a significant difference with the group using detergent only. Two main outcomes; 1) consistency of replicates was minimal with the detergent only antigen but the sonication greatly improved the consistency of replicates (CVs of 30% with detergent vs 6% with sonication) 2) the overall ODs were much greater with the sonicated virus even at quite high dilutions, supporting the notion of greater viral release through sonication. From these results it was evident that the improved difference was brought about by sonication. Sonication has the advantage of liberating more viral protein, and retaining its quality and ability to coat plates.

ii) Secondly, plates were previously left at room temperature overnight to allow the antigen to bind to plates. It has since been established that optimal antigen binding can be achieved at 37°C for 4 hours.

No appreciable difference was detected after 4 hours. This reduction in incubation time results in more efficient coating of plates an aspect that will prove to be important considering the numerous OF samples that will need processing.

iii) Formerly acetone was added to fix antigen onto plates (as used for malaria slides) to facilitate transportation of coated plates from the UK. Until recently lysate was grown in the UK. A culture laboratory has now been set up in Kilifi and the RSV lysate is grown locally thus it is no longer necessary to fix the antigen on plates using acetone. To my knowledge there is no obvious detrimental effect of acetone except that it is no longer necessary.

iv) Previous results suggested that there might be some non-specific binding of general IgG to the plate. This would then mask any true anti-RSV antibody signal. This prompted the need for revisiting the blocking step. The blocking process acts to cover any exposed surfaces on the plate with another protein that does not react with sample antibody thus preventing non-specific binding to the solid phase. Milk (e.g. Marvel powder) an alternative protein that does not react with the sample is usually used and binds to places where Ag is not bound. In the previous protocol milk added to PBS-T was used. It has been suggested that the addition of Tween (a detergent) might hinder the effectiveness of the blocking. PBS with Tween is ordinarily used in plate washing to remove non-specifically bound particles on the plate but not to prevent non-specific binding. Thus is not necessary in the blocking process and has since been left out. In the new protocol plain PBS is used to

constitute the blocking solution. The specific reasons for high background in previous assays are still not clear as there is no known reason for mock (cells without Ag) to show non-specific binding. One possibility is that there may have been some cross contamination in previous culture or during the transportation of plates from Birmingham to Kenya. Presently background levels are consistently low as is desired. This will be important especially for OF samples which have a lower signal thus requires very low background to detect difference.

v) Lastly, the protein adsorption of the plates in use currently is also higher by adopting plates with a high protein binding specification. Standard binding plates appears to result in scant attachment of the viral protein to the solid phase such that once the sample is loaded there is increased chance of unsuccessful binding. Specifically with regards to OF it is felt that this would be important given the low quantity of antibody present in the sample. With a saturated solid phase the chances of successful antigen-antibody binding are higher.

Reluctantly the work on perfecting the oral-fluid method had to be curtailed. The serological element of this project was a key component of the study design insofar as was necessary in defining the contribution of re-infections (age-stratified incidence rates) in both children and older individuals to community transmission dynamics. The absence of these data means that a major aspect of the study has been compromised and results need to be viewed in the light of this omission. However, plans are underway to use this new protocol and assay optimisation with OF samples. If this proves successful we intend to proceed with the processing of all the OF samples collected within the Household study. Current results using the new protocol are shown in Figure 1. This shows the results of a standard checkerboard analysis, varying both antigen and antibody. The results show a clear shouldering effect. These preliminary results indicate that it will be necessary to use a

higher dilution of Ag to coat plates compared with that of there equivalent serum assay. It appears that the 1/4 Ag dilution and neat saliva will be optimal for use; this now requires the culture of large quantities of RSV antigen which is underway

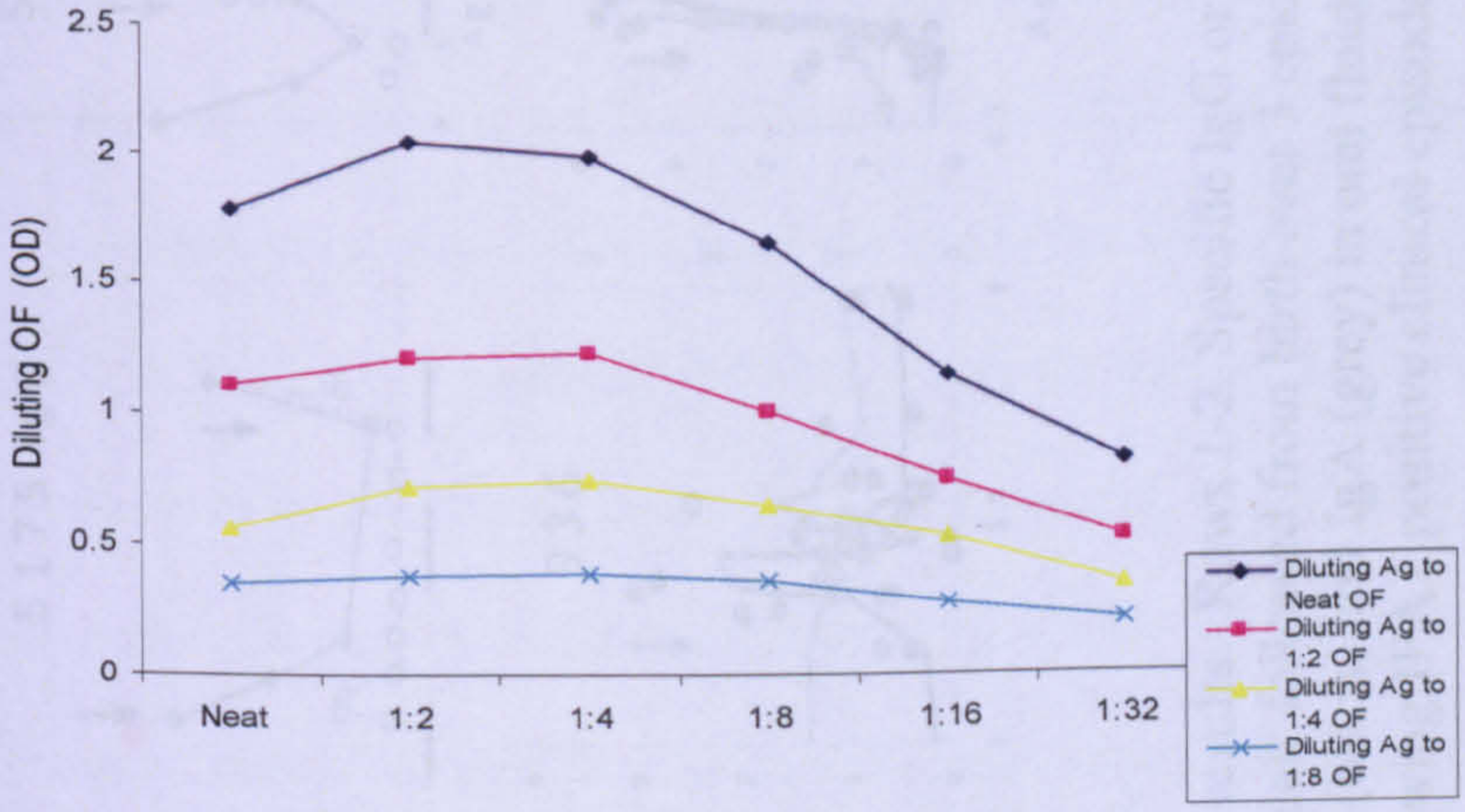


Figure 4.13. Oral fluid checker board titration. Correct OD reading (Top panel). Bottom panel shows actual Ag and mock signals.

The preliminary results of the oral fluid evaluation that have just been concluded are presented illustrating implementation and progress achieved at present.

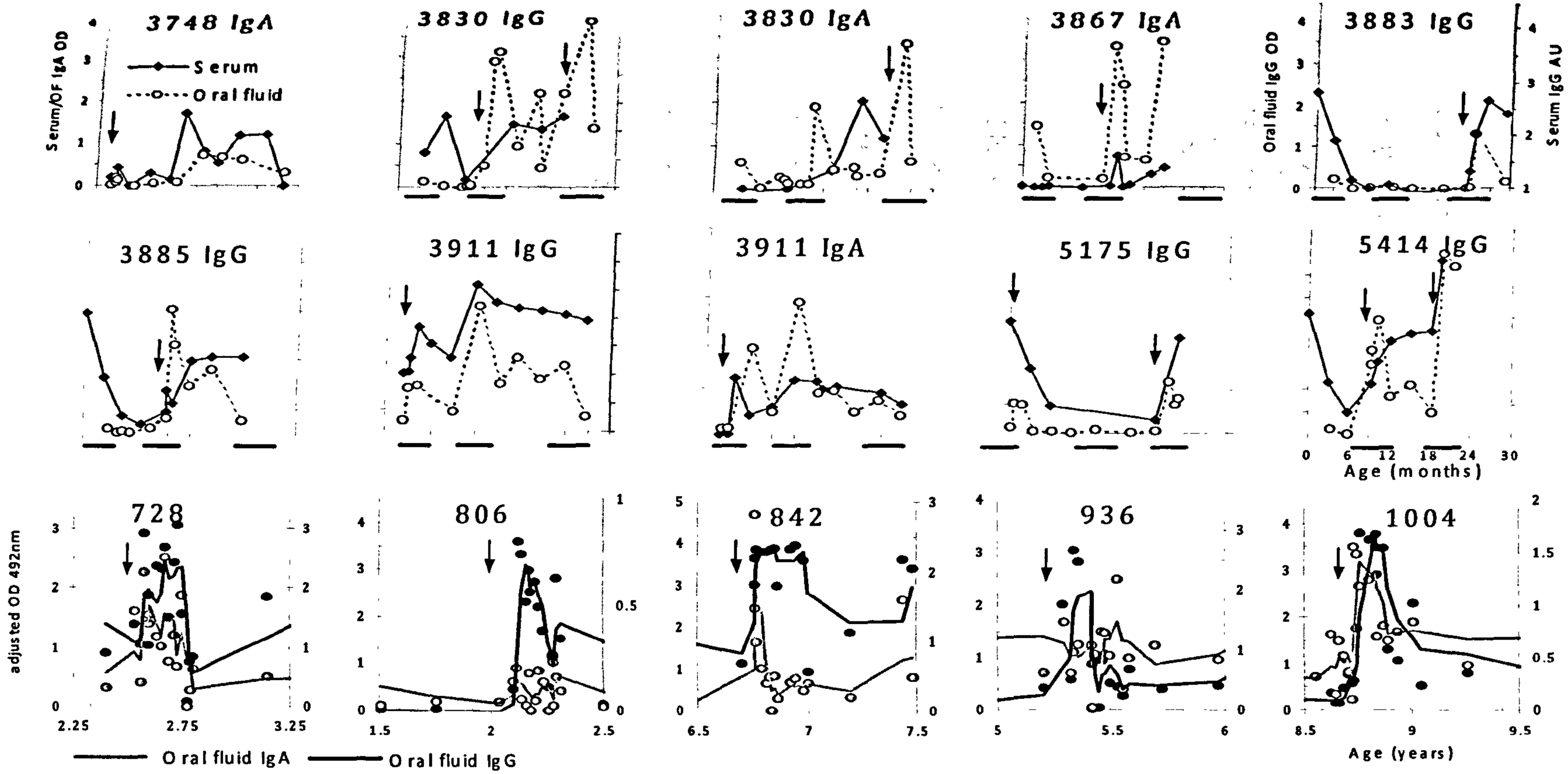


Figure 4.14. Evaluation of oral fluid for the detection of RSV infection: preliminary results. Rows 1-2. Specific IgG or IgA concentrations in serum (filled diamonds) and oral fluid (open circles) samples(2002-5) from 8 individuals followed from birth over 3 epidemics (black bars). Arrows indicate time points of IFAT positive clinical episodes. Row 3. Specific IgG (black) and IgA (grey) in oral fluids from 5 children in a family cohort study sampled every 3 months routinely and weekly for 13 weeks following IFAT positive clinical episodes (arrows). Lines through data are 4 point moving averages.

Chapter Five

Transmission within the household

5.1 Introduction

RSV transmission is known to occur rapidly within a population such that most (60% or more) children are infected during the first epidemic to which they are exposed and practically all have had at least one infection by the end of the third year of life. It therefore seems non-intuitive that RSV is transmitted by large droplets and contaminated materials (fomites) as opposed to aerosol transmission [55, 77] suggesting the need for close contact for infection to occur. Furthermore although most of the disease is observed in young children (< 1 year) it seems unlikely that they are the main source of infection. This hypothesis is supported by evidence from previous community studies (Chapter 2 section 2.7) and some adult studies [53, 98, 135, 136, 142] that report high incidence of re-infections and identify siblings, especially of school age, as an important risk factor (Chapter 2 section 2.8 & 2.11). Accordingly, susceptible individuals in the population can be classified as being of two types: fully susceptible (naïve) children who have not had a primary infection and partially susceptible individuals i.e. those who have partial acquired immunity against RSV and can be re-infected (Figure 5.1.)

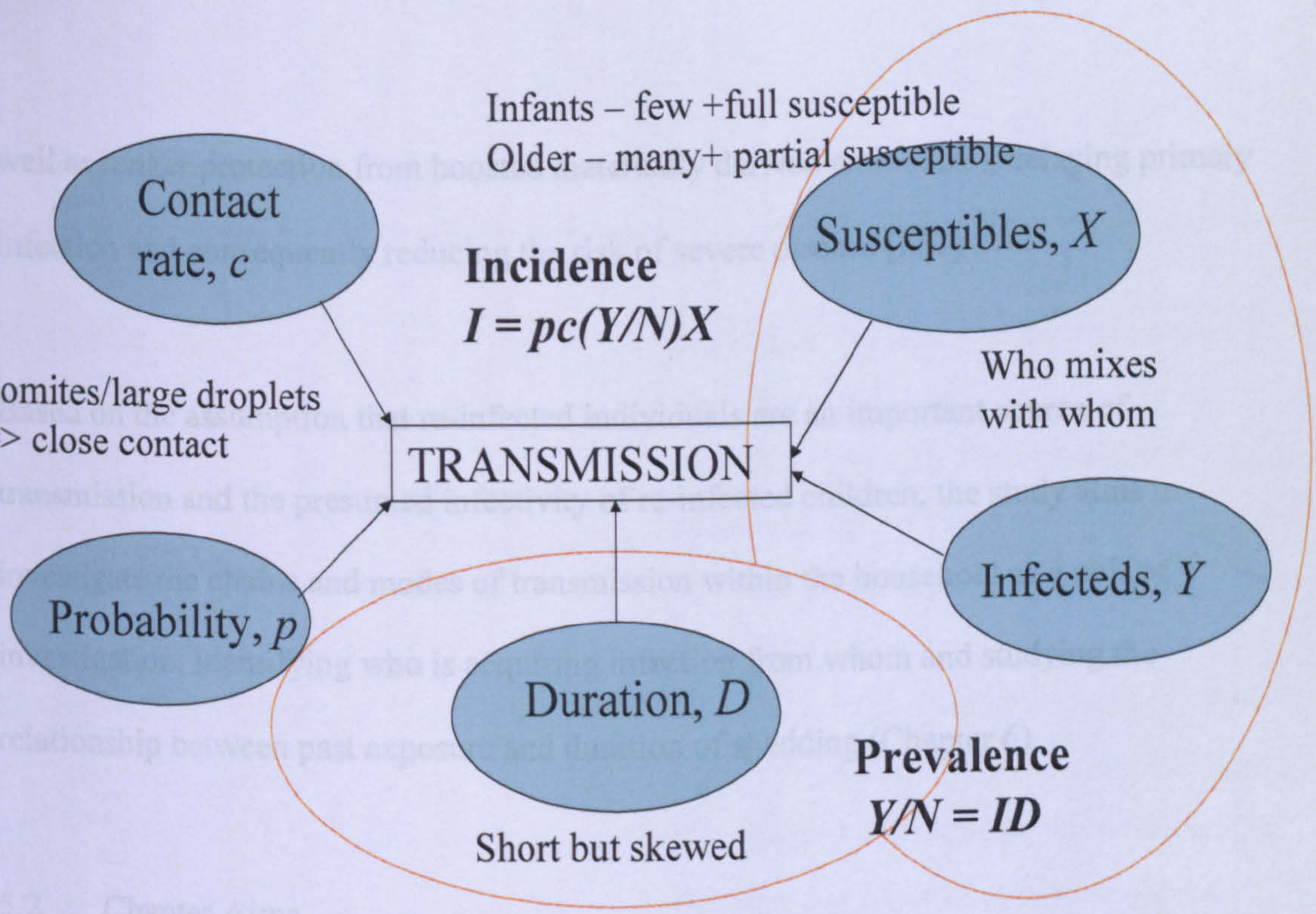


Figure 5.1. Transmission model of RSV. p is the probability of transmission taking place on contact, D is the duration of viral shedding, Y refers to infectious individuals in the population. There are two classes of susceptible individuals (X); fully susceptible infants and partially susceptible older children and adults. Incidence, I is a function of the number (density) of susceptible and the rate of infection per susceptible, also known as the force of infection. The force of infection is a function of the rate of infectious contacts (c Y/N) scaled by the probability of transmission, p. Prevalence is a function of incidence, I, and average duration of infection, D (the inverse of the average rate of recovery).

The epidemiological pattern as represented in the transmission model shows that close contact such as is evident in the home or school setting provides ample opportunity for spread and increased the risk of infection. This is in line with existing risk factors for fomite and close contact spread which include the proximity of child to mother and nearness between siblings in the home [113, 144, 194, 198, 257]. It seems important therefore to understand transmission patterns within the household as a means to understanding transmission at the population level. Examination of levels of mother to child transmission would also be a significant contribution to studies on the possibility of maternal vaccination as a means of reducing the risk of mother-to-child transmission (as

well as longer protection from boosted maternally derived antibodies), delaying primary infection and consequently reducing the risk of severe disease [257] .

Based on the assumption that re-infected individuals are an important source of transmission and the presumed infectivity of re-infected children, the study aims to investigate the chains and modes of transmission within the household as a unit of investigation, identifying who is acquiring infection from whom and studying the relationship between past exposure and duration of shedding (Chapter 6).

5.2 Chapter Aims

This chapter will present the observational data and subsequent chapters will present analyses e.g. Chapter 6 estimate rates of shedding in relation to various risk factors and Chapter 7 estimate rates of transmission. The studies presented in this chapter mainly aim to identify the potential role of re-infection to RSV spread through investigation of re-infection patterns by age, family structure and environmental risk factors; relative contribution of different classes of individuals to transmission; secondary attack rate and factors influencing secondary transmission.

Table 5.1. Table of definitions

Term	Definition
Household	Individuals in a homestead who normally share the same meal.
Sibling	Any child less than 15 years old at the time of recruitment living in the household of the birth cohort child (may include cousins and step-siblings).
Household index case	The first case in a household and the first person to introduce the virus into the home. If more than one case appeared on the first calendar day of an introduction episode, both cases were considered as primary cases.
Household contacts	Persons living in the same household exposed to the primary case.

5.3 Methods

The method of intensive surveillance of a birth cohort and a household cohort was previously described in Chapter 3 and published [5]. In brief, study individuals were monitored through active household visits undertaken weekly during epidemic periods and every 4 weeks otherwise, and via passive referral to a research out-patient clinic or the paediatric wards at the District hospital, which operated throughout the study. The severity of respiratory disease was assigned on the basis of WHO criteria as described previously [5]. The definition of an epidemic was described previously in Chapter 3 with the gap between epidemics defined as the time interval between the end of one epidemic and the start of the next epidemic.

The interval between successive clinical cases in a household (Fig.5.2) was calculated using the method defined by Fine [123]. Briefly, let us assume person 1 transmits infection to person 2, both becoming symptomatic. The observations made are of the clinical onsets and not actual infection transmissions. The separation in time of these cases can be said to be a reflection of two things: i) the timing of infection transmission with regard to person 1's clinical onset (T_1, defined as the time of infection transmission from person 1 *minus* the time of person 1's clinical onset), and ii) the incubation period of person 2 (I_2, being the time of person 2's clinical onset *minus* the time of person 2's infection transmission). Accordingly, T_1 is positive if transmission occurs *after* person 1's illness onset but it is negative if transmission occurs prior to person 1's illness (Figure 5.2). Using this definition, Fine states that: 'the clinical case to case interval between persons 1 and 2 is calculated simply as ($T_1 + I_2$)'. This is a *minimum* if the transmission occurs early in relation (even prior) to person 1's clinical onset (i.e., T_1 is small or negative) and if person 2's incubation period is a minimum (I_2 is small; Figure 5.2, part B). The *maximum* clinical onset interval will occur if transmission occurs late relative to person 1's clinical onset (T_1 is thus a large positive number) and with a maximum incubation period for person 2 (Figure 5.2, part A) [123].

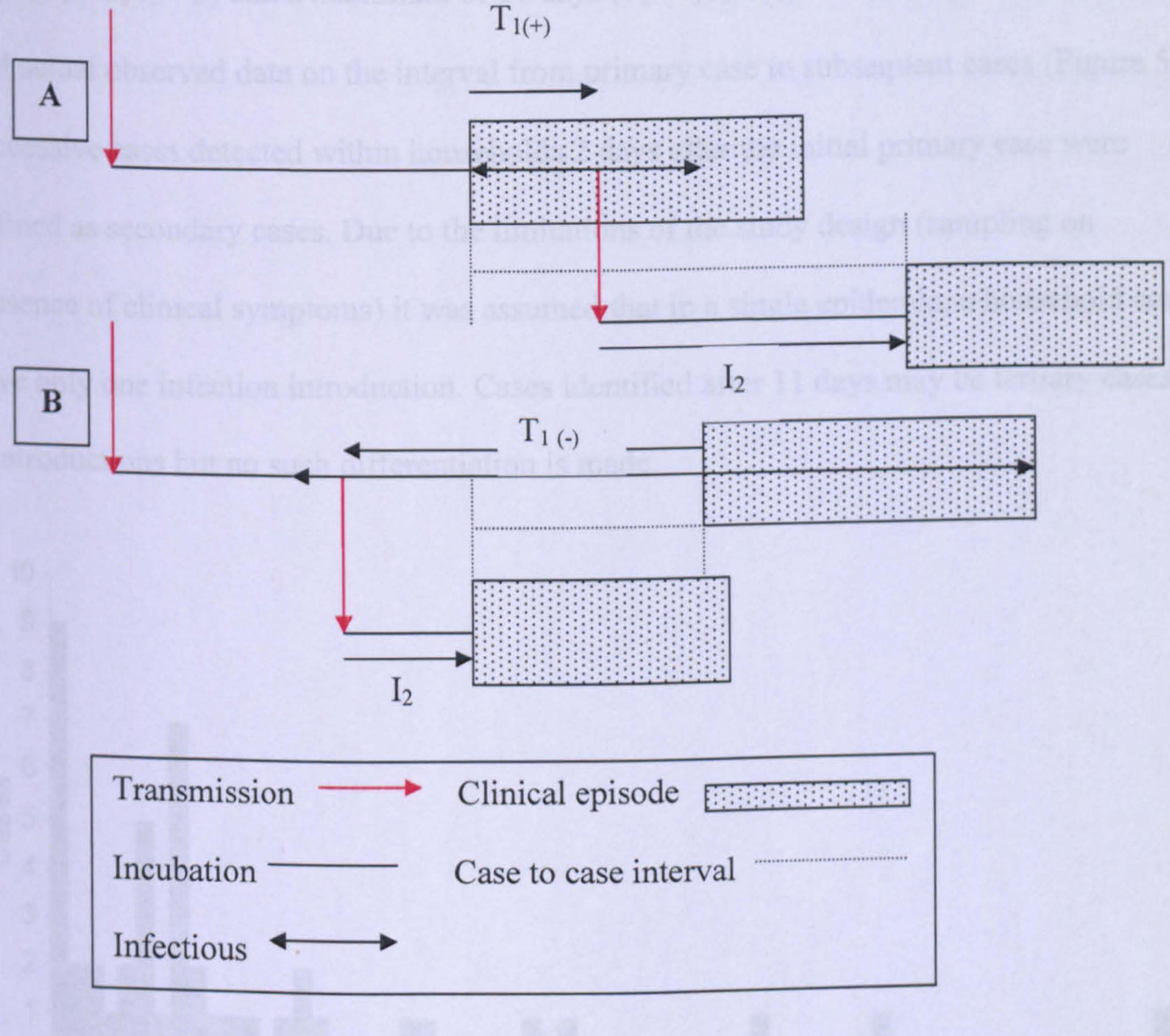


Figure 5.2. Relationship between successive clinical cases (adapted from paper by Fine [123]). In part A, ($T_1>0$) as transmission from person 1 occurs relatively late in the course of 1's infection and case 2's clinical onset occurs (T_1+I_2 time units) after that of case 1. In part B, the combination of early transmission, before case 1's clinical onset ($T_1<0$), and a short incubation period for case B ($I_2< (-T_1)$) means that case B is symptomatic before case 1.

T_1 and I_2 are not directly observable in this study therefore evidence from previous studies [57, 58] discussed in Chapter 2 was used to estimate the two intervals. From this controlled adult challenge study, a preclinical infectious period of 1-3 days and an incubation period of 3-7 days were observed. The range in days from infection to viral shedding was 2-5 days with observed recovery of virus for 3-4 days after clinical onset. The assumption made is that the start of infectious period is correlated to viral shedding (discussed in Chapter 6).

Using this method the case to case interval was estimated to be between a minimum of 0 days (T_1=-3; I_2= 3) and a maximum of 11 days (T_1=4; I_2 = 7). With this estimated interval and actual observed data on the interval from primary case to subsequent cases (Figure 5.3) successive cases detected within households 2 days after the initial primary case were defined as secondary cases. Due to the limitations of the study design (sampling on presence of clinical symptoms) it was assumed that in a single epidemic, a household could have only one infection introduction. Cases identified after 11 days may be tertiary cases or reintroductions but no such differentiation is made.

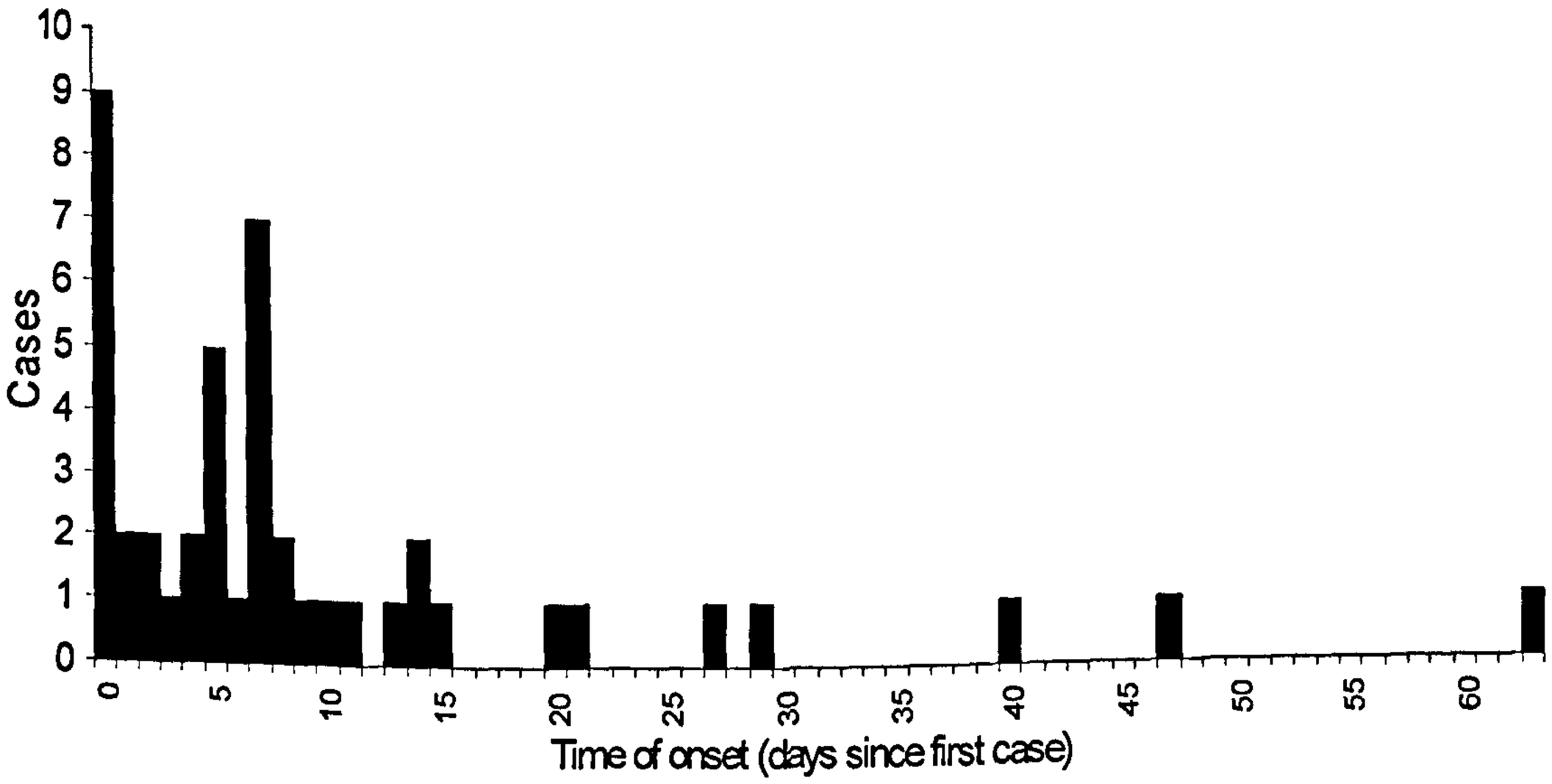


Figure 5.3. Frequency distributions of the interval from primary case onset to the onset of subsequent cases in households within the space of one epidemic.

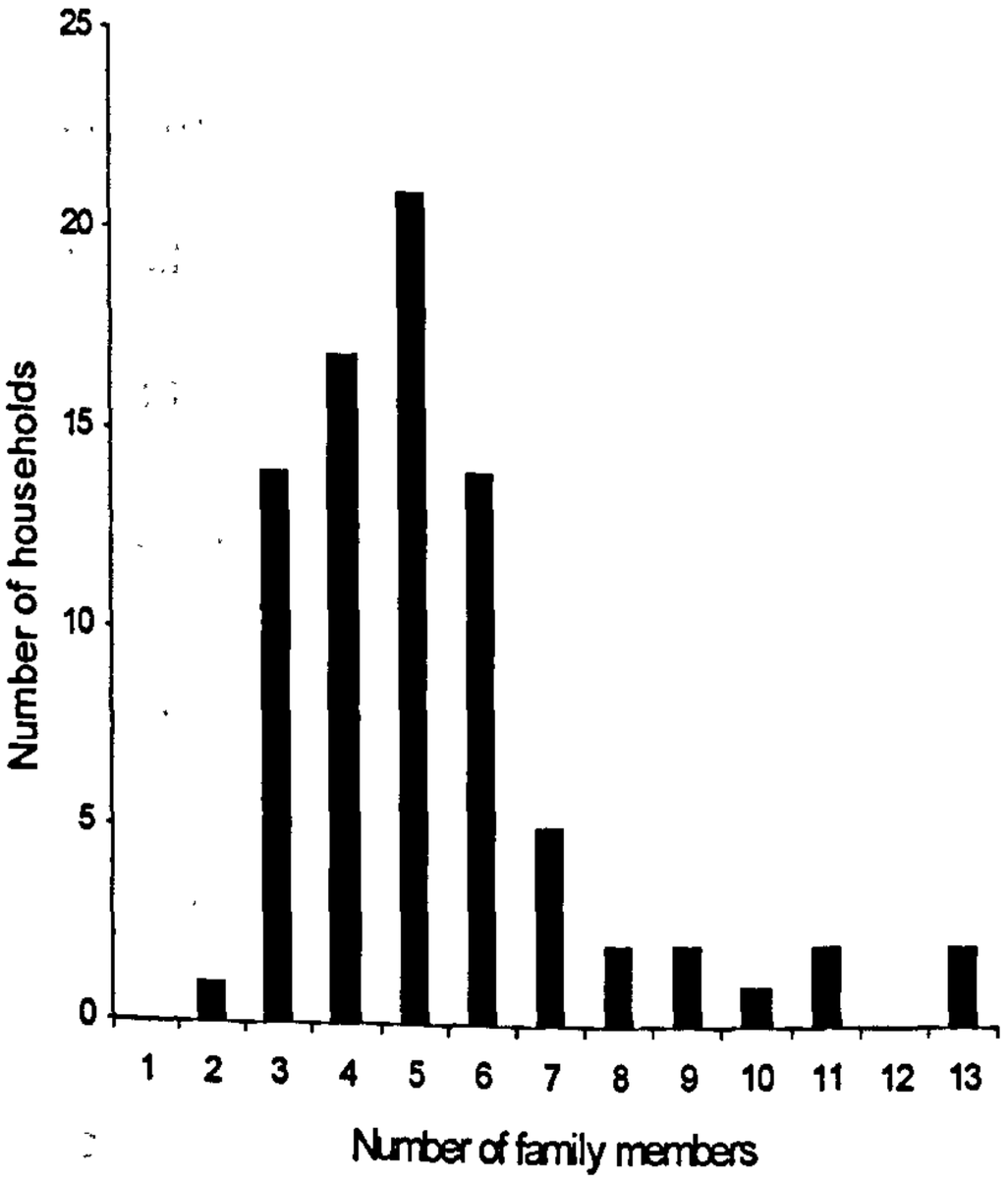
5.4 Results

Details on numbers of households and individuals recruited and losses to follow up (by year or epidemic) have been provided in the results section of Chapter 3.

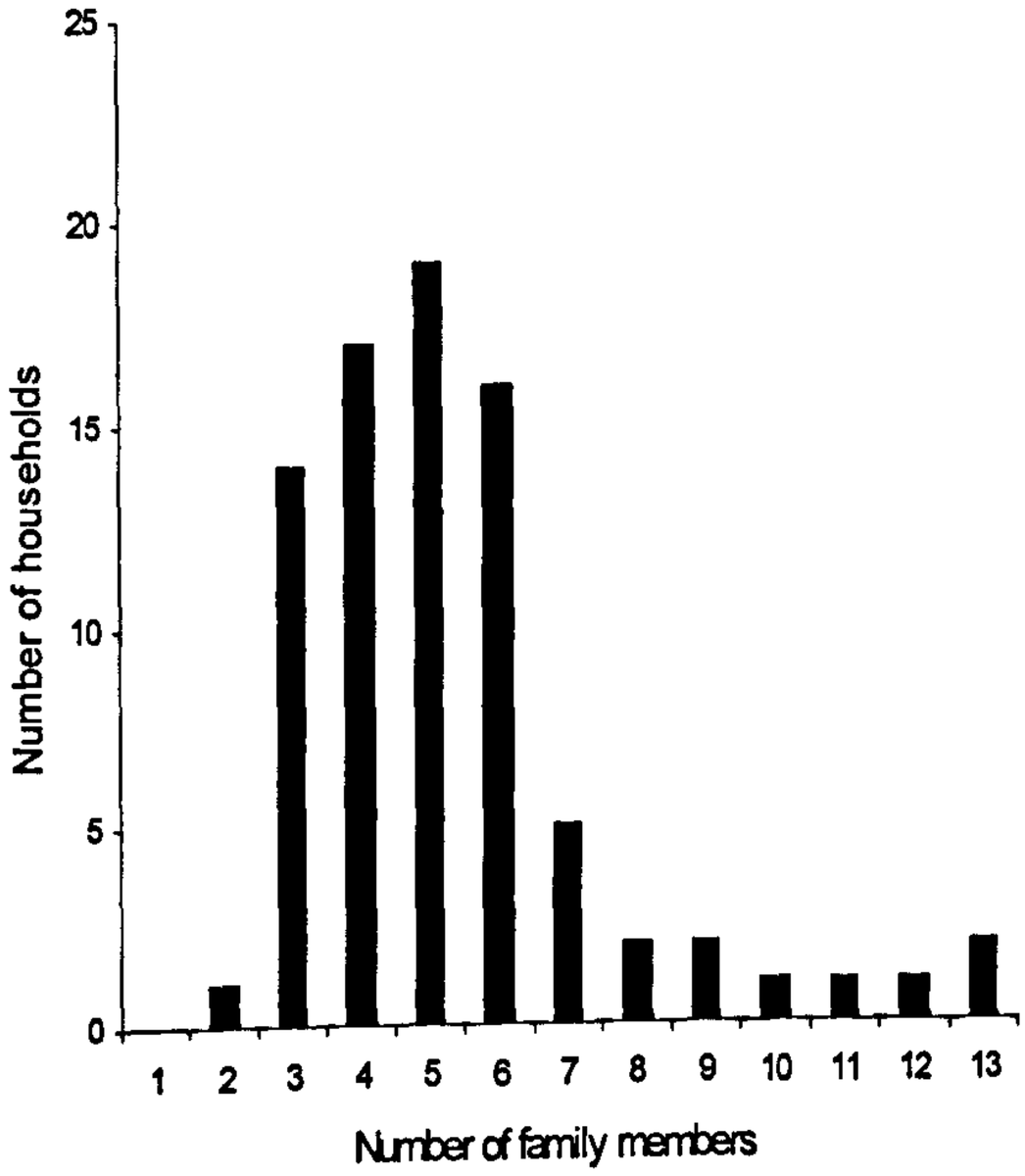
i) Description of Households

Families in this study had between 2 to 13 members, Figure 5.2 with a median of 7 members and an average of 5.3 members. During the follow up period some households

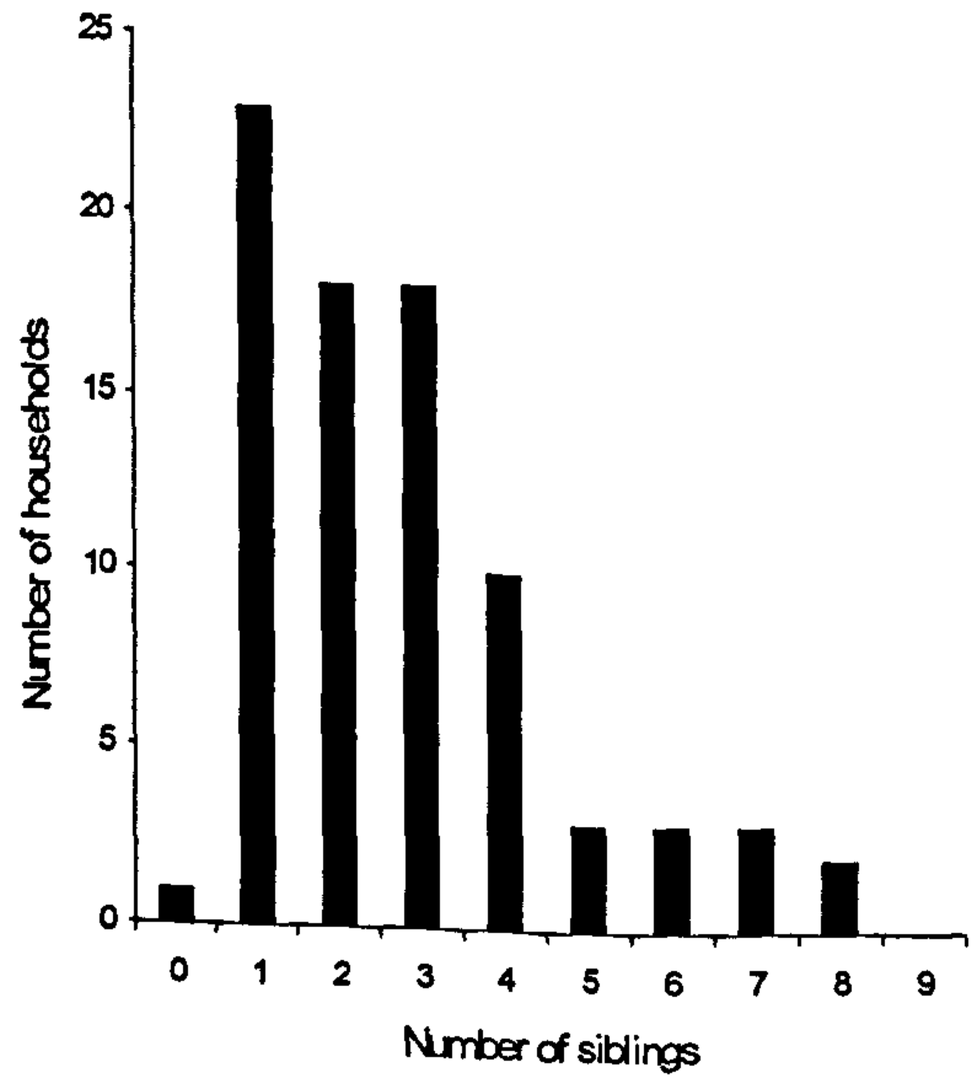
had new members introduced through births, Figure 5.2a (start of the study) versus b (end of the study). Individuals who were under 15 years of age at the time of recruitment were classified as children and are herein referred to as "siblings" to the birth cohort child. The study households had between 0 and 9 siblings as shown in figure 5.2 c (start of the study) and d (end of the study) with an average of 2.8 siblings. The birth cohort child was the only child in one replacement household.



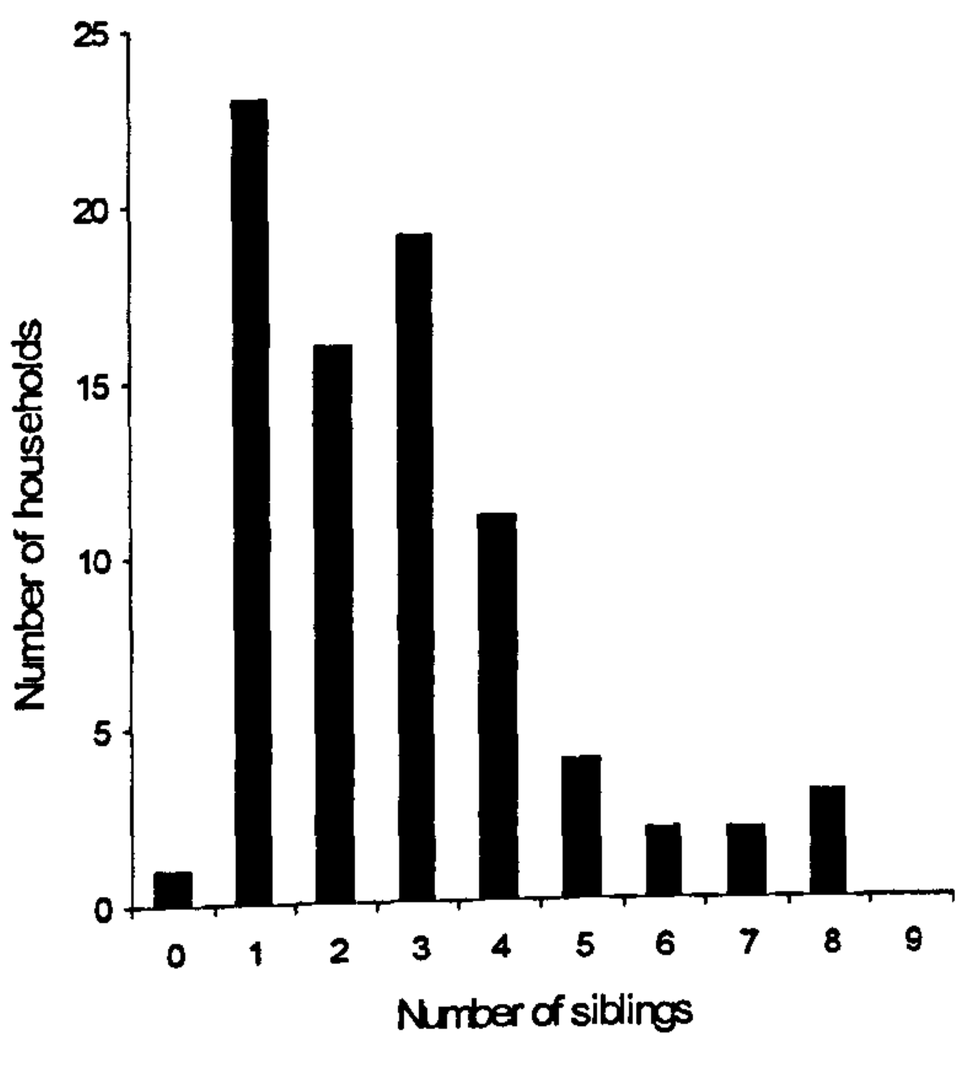
a.



b.



c.



d.

Figure 5.4. Characteristics of households at the beginning and at the end of the study. a) Household sizes at start of study, b) household size at end of study, c) number of siblings in the study households at start, and d) number of siblings at the end of study.

ii) Age distribution of participants

At recruitment siblings of the birth cohort child ranged in age between 0 and 176 months (14.6 years). 55.9% of the siblings were five years of age or younger (Figure 5.3 a). The adults in the study were between 15 and 81 years of age with 80% being less than 35 years of age. 10% of the adults were less than 20 years of age (Figure 5.4 b).

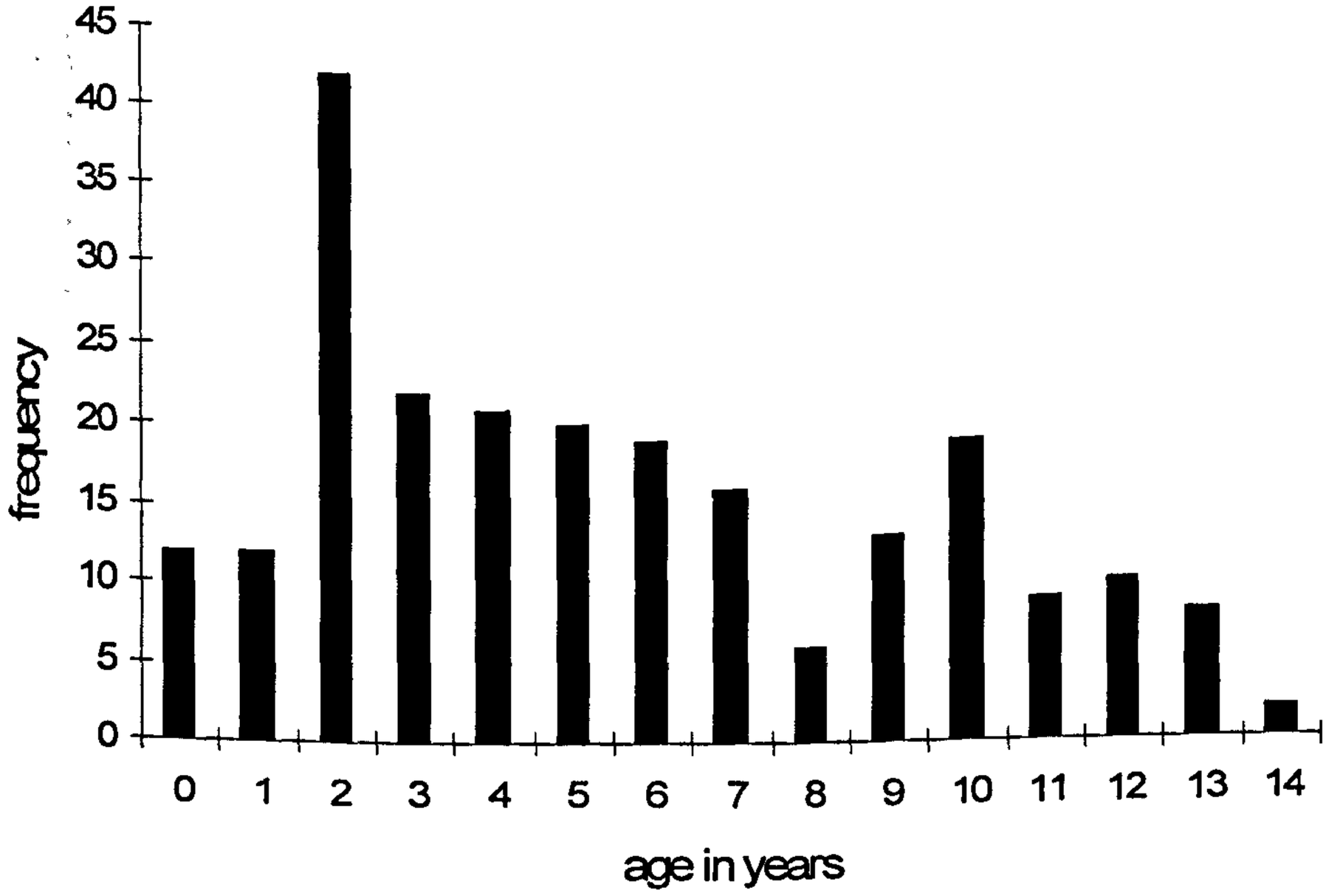


Figure 5.5a. The age distribution of siblings in the household study

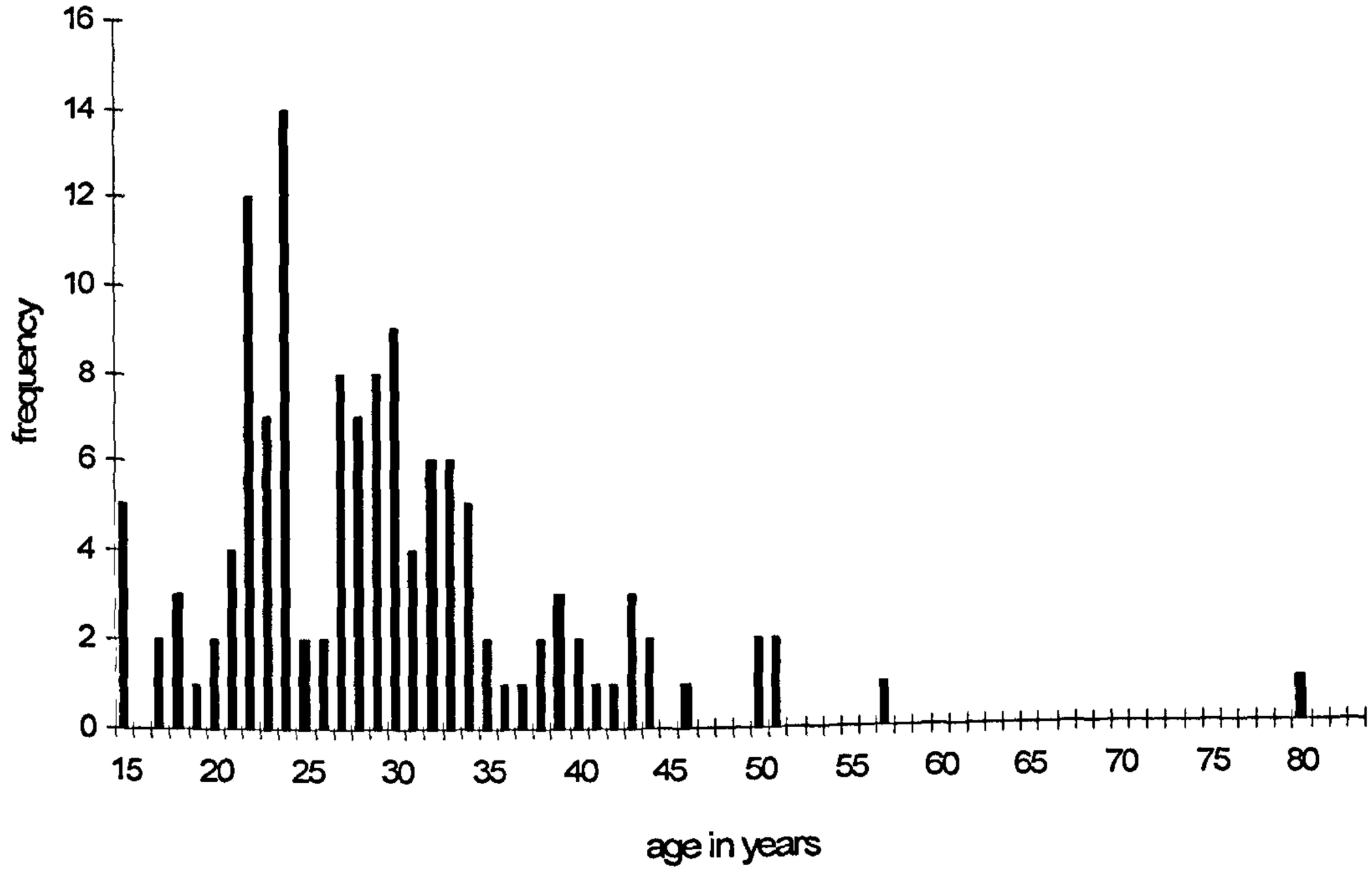


Figure 5.5b. The age distribution of the adults in the household study

iii) Sex distribution

Overall there were slightly more females than males in the household cohort. Within the child (sibling) and birth cohort categories there were equal numbers of males and females but there were considerably more adult females 75% than adult males 25 % in the study as shown in Table 5.2, consistent with the demography of the District (Figure 3.2).

Table 5.2. Sex distribution of household individuals

	Cohort type			
Sex	**Adult**	**Child**	**Birth**	**Total**
Female	99	115	43	258
%	75.0	49.8	51.2	57.6
Male	33	116	41	190
%	25.0	50.2	48.8	42.4
Total	**132**	**231**	**84**	**447**

iii) Epidemics

The household cohort experienced three epidemics in the population, defined as epidemic 1-3, within the dates 3/12/02 to 15/4/03, 8/1/04 to 2/6/04 and 11/11/04 to 18/2/05, respectively. The interval between the successive epidemics (i.e. end of one and start of next) was 268 days (8.8 months) and 162 days (5.3 months), respectively. Recruitment of households started in the middle of epidemic one. 73 households were recruited in the first epidemic, with 400 participants. The denominator (number of households) at each successive epidemic was 67 and 62 households respectively with 367 and 353 study participants, respectively.

5.4.1 Infection and re-infection patterns

A total of 121 clinical infections were identified, and occurred in 53 out of 81 (54%) households. One replacement household is excluded from analysis involving introduction of infection into the household and transmissibility because the birth cohort child became infected before the household was recruited. The distribution of individual infections and household infections by weeks during follow-up is shown in Figure 5.4.

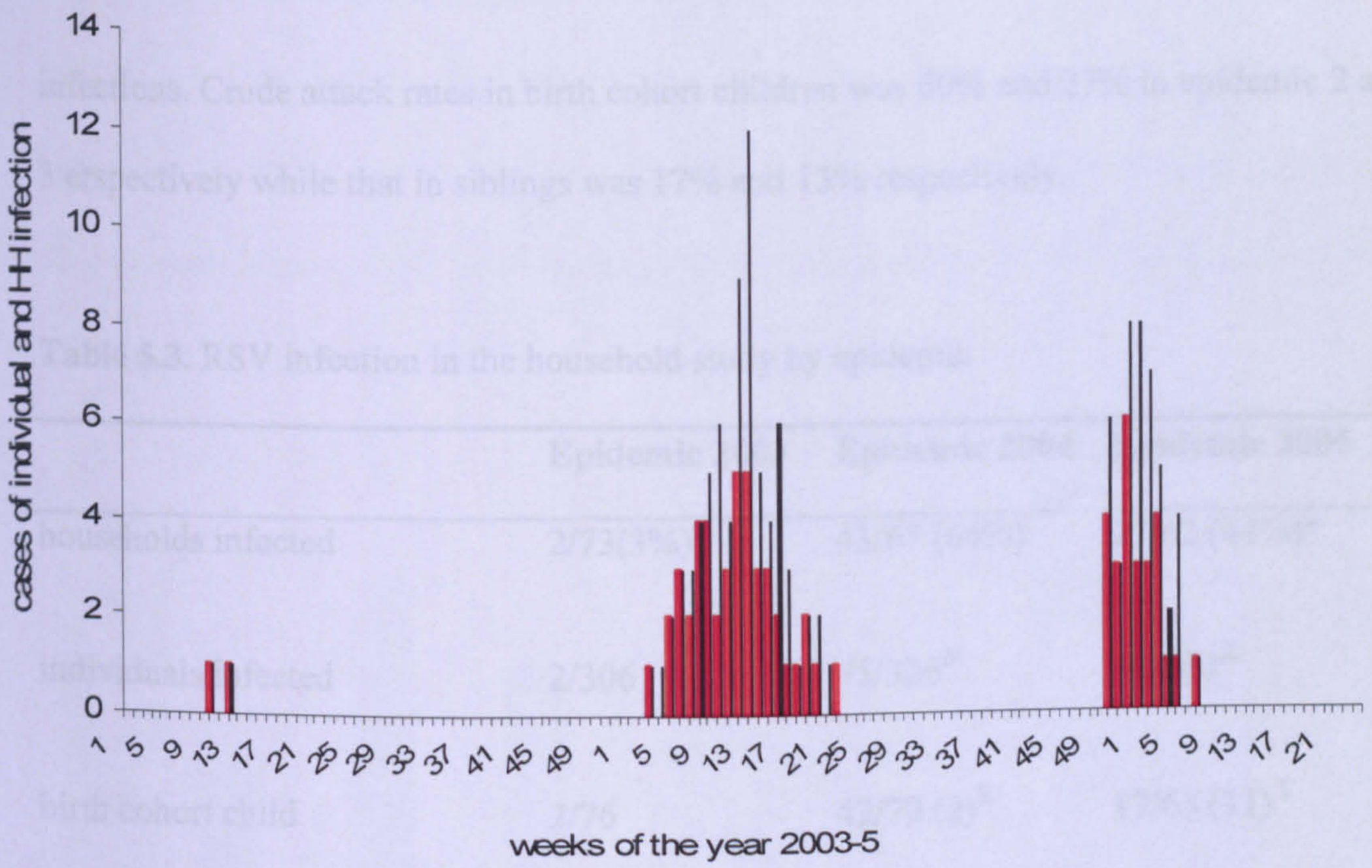


Figure 5.6 Distribution of individual infections (black) and household infections (red) by week

In epidemic one, two birth cohort children from two households were infected. In the second epidemic there were 75 clinical infections identified in 43 households, while in the third epidemic 41 infections were identified in 27 households (Table 5.3). 20 (74%) of the 27 households in epidemic three had previously been infected. Of the 75 infections identified in epidemic two, 42 were in the birth cohort children who were all less than 1.5 years old at the time. In this epidemic, only two re-infections were identified both arising in birth cohort children. In the third epidemic 17 infections were in birth cohort children and of these 11 were re-infections. Out of the 23 infections identified in the siblings, 7 (30%) were re-infections. Of these, 5 were infected in epidemic two while the other two siblings, one 3.8 years old and the other 1.2 years old, were infected twice in the third epidemic (Table 5.3). With the exception of two instances, the interval between the two RSV infections was less than 6 months the interval between epidemic two and three being only 5 months. In those 2 exceptions there were spans of 22 and 23 months between the two

infections. Crude attack rates in birth cohort children was 60% and 27% in epidemic 2 and 3 erspectively while that in siblings was 17% and 13% respectively.

Table 5.3. RSV infection in the household study by epidemic

	Epidemic 2003	Epidemic 2004	Epidemic 2005
households infected	2/73(3%)	43/67 (64%)	27/62 (44%)*
individuals infected	2/306	75/326[&]	41/300[&]
birth cohort child	2/76	42/70 (2)[$]	17/63 (11)[$]
siblings(0-15 years at recruitment)	0/230	33/199 (0)[$]	23/182 (6)[$]
adults	0	0/57	1/55

* 7 new introductions
$ (re-infected)
& denominator includes adults under clinical surveillance

43% (36/84) of birth cohort children were infected on one occasion only during the study compared to 19% (43/230) of siblings. 15% (13/84) of birth cohort children were infected on more than one occasion, 2 of whom were infected three times, compared with 3% (6/230) of siblings with re-infection, 1 of whom was infected 3 times. Hence in total 19/314 (6%) children were re-infected with a total of 22 re-infections.

i) Infection and re-infection patterns in relation to age

The age stratification of all the infections identified during the household study is shown in Figure 5.5. The majority, 88%, of infections were in children less than 5 years of age. 67%

of the infections in children <5 years occurred in children 3 years of age or less. 13 (11%) infections were identified in children between the ages of 5 and less than 10. 73% of reinfections identified during the study occurred in children ≤3 years, 68% (15/22) were in birth cohort children (i.e. <3 years), and 32% in siblings. 86% of sibling infections were in the age group 3 -<5 years.

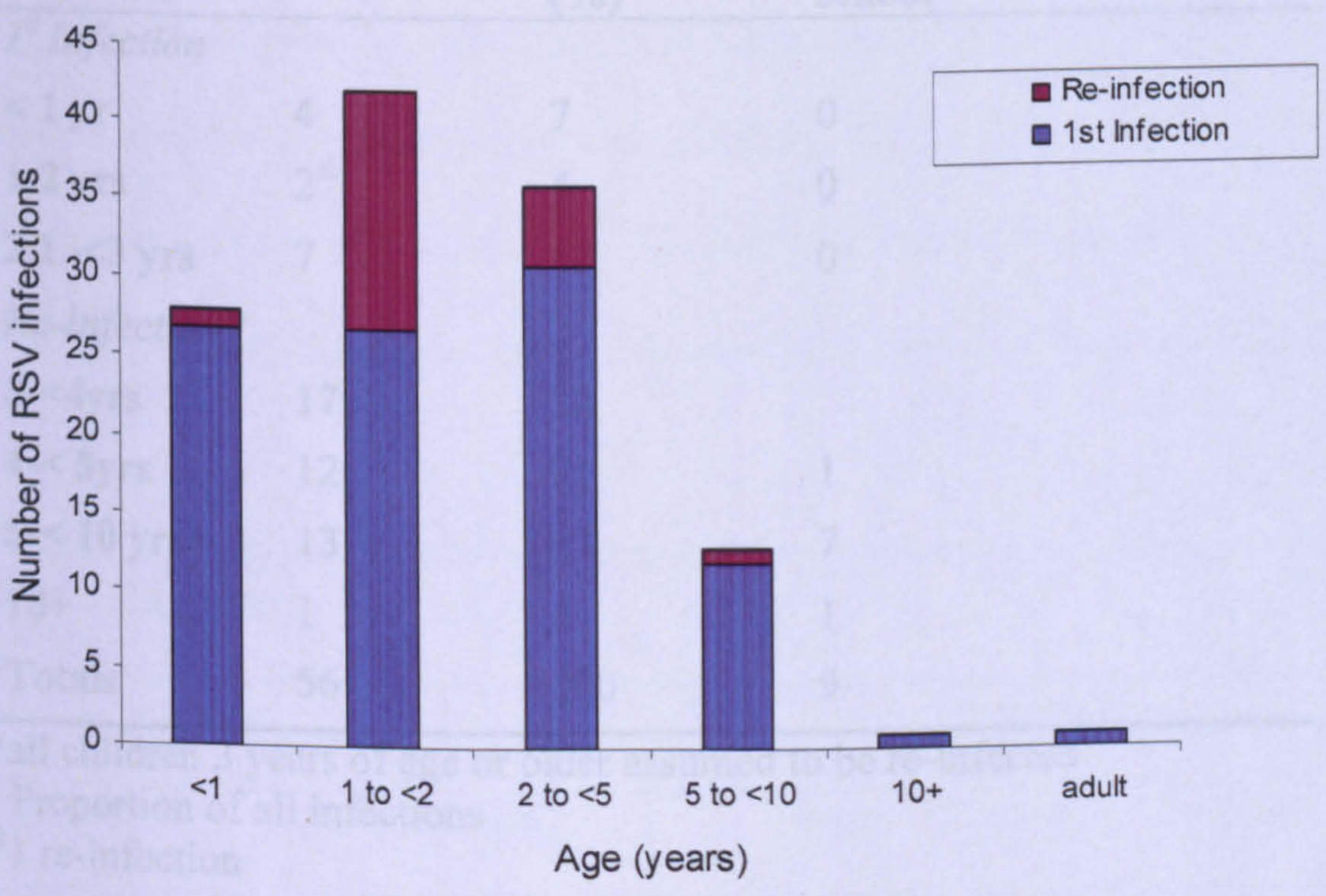


Figure 5.7. First infections (blue) and re-infections (maroon) in relation to age in a household cohort. Adult is defined as individuals ≥ 15 years at recruitment

The data shown in Fig 5.7 are stratified as first and re-infection on the basis of observation during this study. However, it may be inferred from previous studies (see section 2.4 Chapter 2) that all siblings who were 3 years of age or older were in fact most likely to be suffering from a re-infection (Table 5.4). This assumption is adopted for all subsequent data analysis in this Chapter. Thus, 76.8% of the infections observed in siblings were re-defined as re-infections. 75% of these re-infections were in children between 3 and less than 10 years of age. 10.7% of the infections in siblings were in household children younger than

the birth cohort child (born during the study). Of the infected siblings, only 16% (9/56) were attending school. The age-specific incidence rates are described in Chapter 7.

Table 5.4. Infections in siblings in the household cohort

Age (years)	Number	Proportion[$] (%)	Number attending school
1⁰ Infection			
< 1 yr	4	7	0
1-2 yrs	2[&]	4	0
2.1 -<3 yrs	7	13	0
Re-infection*			
3-<4yrs	17	30	
4-< 5yrs	12	21	1
5-< 10 yrs	13	23	7
10+	1	2	1
Totals	56	100	9

*all children 3 years of age or older assumed to be re-infected
$ Proportion of all infections
&1 re-infection

ii) Infection and re-infection patterns in relation to family structure

A comparison was made between households in which infection occurred, 52 (65%), and those in which no infection occurred, 28(35%), to identify household characteristics that may influence observed infection patterns. 8 households were lost to follow up before experiencing an epidemic and thus are excluded from this analysis. The number of children (<15 years), household size or number of adults in the home did not differ between the infected and non-infected households (χ^2 test, $P > 0.05$). Similarly infected and non-infected households did not differ with regard to the classes of children present in the home: i) number of pre-school children ii) number of school children under 6 years and iii) number of school children 6 years of age or more (χ^2, $P > 0.05$).

Comparison based on size of outbreak

Data on the size of outbreaks (number of cases in a household) within different households are presented in Table 5.5. In examining if there are certain household and environmental factors that promote the transmission of RSV, several factors were tested. Households are classified according to both type of households and the number of cases. A comparison of the outbreak size for the different types of households is equivalent to testing the independence of the two types of classifications. Of the factors tested, only socio-economic status was found to affect the size of the outbreak in households ($\chi^2(2) = 10.6$, P= 0.005). Further analysis of individuals and household risk factors is reported in Chapter 8 and 9.

Table 5.5. Factors that influence outbreak size (number of infections in a household) in the household cohort.

		Outbreak size				
Factors	**Categories**	**1 case**	**2 cases**	**3+ cases**	**Totals**	**P-value***
household size	10+	4	2	0	6	0.099
	5 to 9	22	19	7	48	
	<5	15	4	0	19	
SES$	poor	25	23	7	55	0.005
	not poor	16	2	0	18	
pollution	burn refuse	34	22	4	60	0.166
	don't burn	7	3	3	13	
cooking location@	different house	22	12	4	38	0.678
	same house	13	11	3	27	
	outside	6	2	0	8	
fuel	firewood	32	21	7	60	0.358
	other&	9	5	0	13	
sanitation	no toilet	15	7	4	26	0.356
	latrine or flush toilet	26	18	3	47	

* χ^2 test

$ Socio-economic status (defined using asset index score (Chapter 8)). Households sorted by asset index and establish cutoff values for percentiles of the population.

@ House location same as or different from sleeping area

& Other - (paraffin/charcoal/mix)

5.4.2 Introduction of infections into families

There was a suggested association between introduction of infection into the household and household size (Figure 5.8) with greater proportion of larger households infected. Borderline significant association between HH size and occurrence of an outbreak observed above (χ^2, P=0.09).

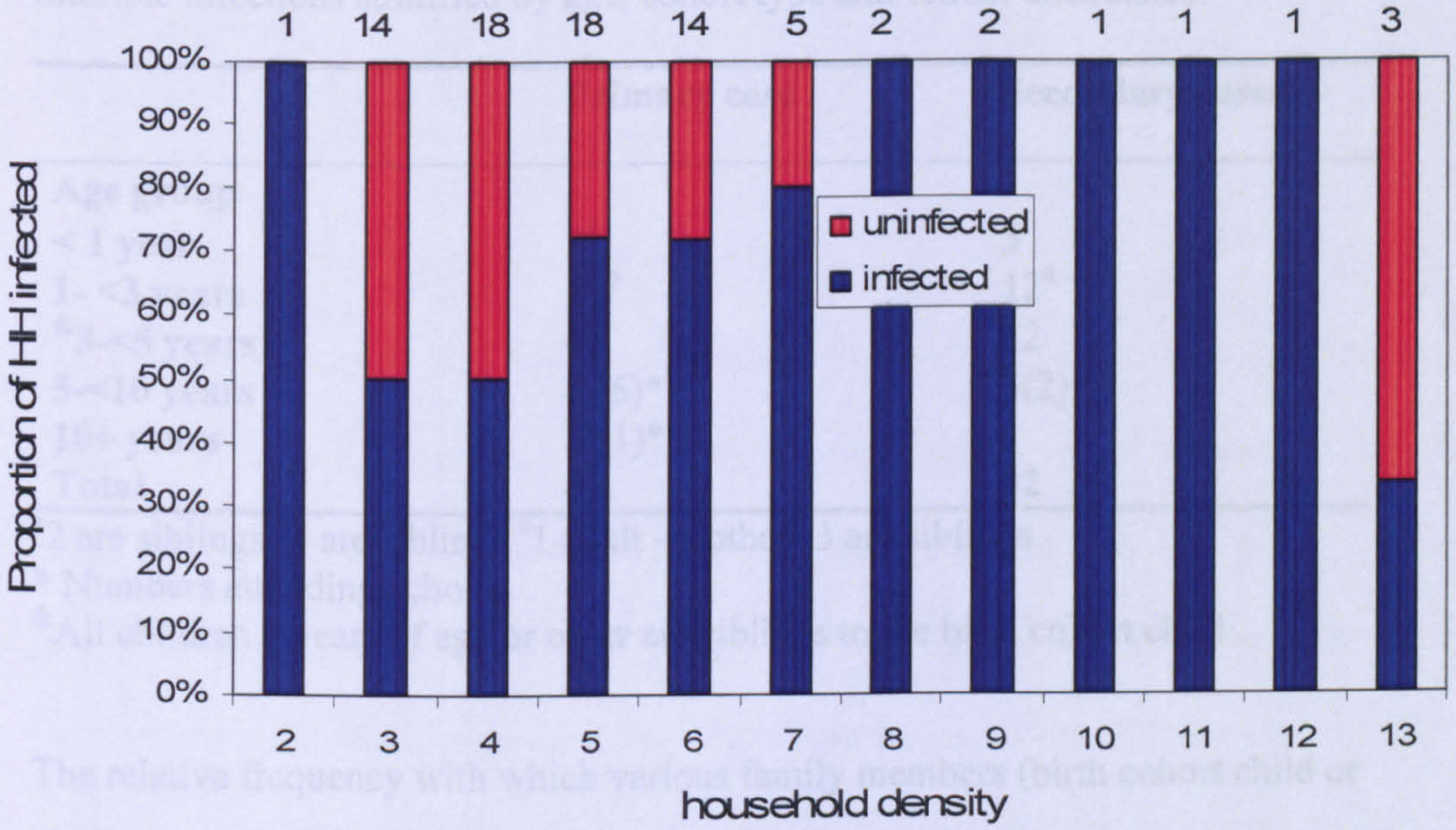


Figure 5.8 Proportion of households into which virus introduced. Numbers above each bar show frequency of households of each size.

Single primary cases occurred in 38 families in the three epidemics and co-primary cases in an additional 8 families. Within co-primary cases, two of the 8 cases included an infant. Secondary transmission occurred in 25 instances in two epidemics, 17 households in epidemic two with 5 having more than two household members infected and 8 in epidemic three. Table 5.6 shows the primary and subsequent case classification by age for epidemic two and three. 78% of primary cases were of pre-school age (< 5 years). Only 7 sibling primary cases were school attendees. Only 1 case of an adult introduction was documented out of the 41 households in which adults were also under routine clinical surveillance from

the middle of epidemic two. In 6 of these households the father worked and lived away from the home and was therefore not likely to introduce infection or be infected by other household members.

Table 5.6. Characteristics of primary and secondary cases in RSV positive families with multiple infections stratified by age, cohort type and school attendance.

	Primary case	**Secondary cases**
Age group		
< 1 year	8[a]	5
1- <3 years	17[b]	12[d]
[&]3-<5 years	12	12
5-<10 years	8 (6)*	3(2)*
10+ years	2[c](1)*	-
Total	**47**	**32**

[a]2 are siblings [b]4 are siblings [c]1 adult –mother [d]3 are siblings
* Numbers attending school
[&]All children 3 years of age or older are siblings to the birth cohort child

The relative frequency with which various family members (birth cohort child or sibling/adult) introduce RSV infection into the home was investigated [258]. Birth cohort children were most often the index case identified (Table 5.7); except in families of 6-9 members where the siblings were more often identified as the index case. When households with a single case of infection were excluded, siblings were 1.4 times more likely to introduce the virus into the home.

Table 5.7 Relative frequency of introduction in households of different characteristics

Household Characteristics		Relative risk of introduction*	
HH size	Number of children in HH	No.	RR*
2-3 members	1-2	9/2	4.5
4-5 members	2-4	17/8	2.1
6-9 members	4-7	15/24	0.6
10-13 members	8-9	5/2	2.5

* $\frac{\text{Number of introductions by birth cohort child}}{\text{Number of introductions by siblings/adult}}$

5.4.3 Spread of RSV infection within the home

The frequency of secondary cases that developed on each day after a primary case has been detailed in Figure 5.3. An index of the frequency with which spread occurs, the secondary attack rate (SAR), was defined as the number of secondary cases divided by the number of family children exposed to the index case(s). In Table 5.8 we first show the relationship between the age of index cases and age of secondary cases. Overall there was little evidence to suggest any relationship between age of index and age of secondary case. However, older primary cases (5-< 10 years) were observed to infect younger children more often than older children while the reverse was true for younger index case (<3 years). It is expected that this result would particularly be evident in smaller HHs since the first case 'uses up' one of the two groups (younger or older). Due to small numbers no such subgroup analysis was done.

Table 5.8 Matrix of age of primary case and age of secondary cases

	Age of secondary case (years)				
Age of index case (years)	**<1 (%)***	**1-<3(%)**	**3-<5 (%)**	**5-<10(%)**	**Total**
<1	0(0)	3 (43)	3 (43)	1 (14)	7
1-<3	2 (20)	0(0)	8 (80)	0(0)	10
3-<5	0(0)	6(86)	1(14)	0(0)	7
5-<10	3(43)	3(43)	1(14)	0(0)	7
10+	0(0)	1(33)	0(0)	2(67)	3
Total	**5**	**13**	**13**	**3**	**34**

*Proportion of secondary cases infected by age class of index case

Interval between primary and secondary case

The occurrence of subsequent infections after an introduction was explored in terms of the interval between the index cases and the secondary cases, Table 5.9. Median interval between index and secondary cases was 7 days with a mean of 12.4 days. The majority (59%) of secondary transmission appeared to take place within 8 days of introduction- most transmission taking place between 5-7 days. The youngest age group seemed to experience significantly shorter (t=-1.95, P=0.03) intervals thus appear to be infected preferentially or it may be the case that they have a shorter incubation period. There was only one primary case classified as severe RSV-LRTI with two others as RSV-LRTI. There were 4 cases of disease among secondary case; 2 classified as severe RSV-LRTI and 2 RSV-LRTI (in twins). Thus there were too few cases to assess the relationship between the time to secondary case and the severity of the first case, or the severity of the first case to the severity of the secondary case.

Table 5.9. Median interval between index and secondary cases of RSV infection in family children by age of the secondary case.

age of secondary case (years)	Numbers	median interval (days)	mean interval (days)
<1	5	5	5.4
1-<3 years	12	8	15.1
3-<5 years	12	7.5	11.6
5-<10 years	3	5	16.7
10+	0	no cases	no cases
Totals	32	7	12.4

Secondary attack rate

Secondary attack rates (SAR) appear to decline with increasing age of contacts. The mean SAR tended to be higher in younger children, those less than 3 years, than in older children 3- <5 years (0.48 vs. 0.34, (t=1.87, P=0.07)) and 3-<10 years 0.48 and 0.35 (t=1.81, P=0.08) respectively although not significantly. Households of smaller size had significantly higher SAR. The mean SAR was 0.72 for households with 2 children compared 0.37 and 0.21 for larger households with 3 to 5 or 6 or more children respectively (t=3.34, P =0.004; t=4.70, P=0.002). The mean SAR was significantly higher for households with 3-5 children compared to those with 6 or more (t=2.51, P=0.021). Smaller households had on average younger children. The relationship between age-specific secondary attack rates and the age of the index case (Table 5.10) is presented for family episodes with a single index case of infection. There was evidence to suggest that the secondary attack rates were related to the age of the index case. The mean SAR following exposure to a child under 1 year of age was generally lower than if the index case was an older child (column 2 and 3 vs. 4). Older children (4+ years) tended to be more effective in spreading infection to younger children than to children of the same age (0.67 & 0.63 vs. 0.41) though this was not significant. On average, older index cases seemed to spread

significantly more infection to younger contacts than younger index case (1-<4 years) to children the same age (0.63 vs. 0.35, t=-2.55, P=0.059).

Table 5.10. Age-specific secondary attack rates according to the age of index cases, in episodes with a single index case

	Age of index case (years)		
Age of contact (years)	**<1**	**1-<4**	**4+**
<1	-	0.25	0.67
1-<4	0.36	0.35	0.63
4+	0.33	0.40	0.41

5.5 Discussion

Several aspects of infection and transmission of RSV within the household and results concerning the introduction of infection and secondary transmission are presented. In general birth cohort children were more often the index case identified in the home more frequently than older siblings. In developed countries young children are exposed to a large group of non-household contacts through daycare and other pre-school activities, activities that are alien to this community, hence are more at risk of acquiring infection outside the home. The above result is therefore surprising given that most young children in this rural Kenyan community have limited contact with other community members. The result may well be a reflection of lower immunity in these young children leading to greater probability that infection is symptomatic accompanied by higher likelihood of dectection (increased viral shedding and longer duration of illness [62]). This is an issue explored in Chapter 6. However, when only households with multiple infections (i.e. secondary cases) were considered, siblings were found to introduce infection into these homes more often. This result is similar to reports from other studies [15, 50] and is in line with the

expectation that older children have more potential contact infectious people in the community (e.g. school, and play areas). Nevertheless, only 7 index cases were associated with school attendance (discussed further below). Because of the general mildness of most respiratory infections specifically when they occur in older children and adults [16, 118], an apparent introduction into the family may actually have been a secondary acquisition from an inapparent index case within the home. This possibility of sub-clinical infection to which our study is insensitive could also be an explanation for the near complete absence of (a single case only) of infection identified in adult participants; it is possible that symptomatic re-infection is very uncommon in this setting.

Of the previous family studies carried out in developed countries only one study [15] used symptomatic identification of infection -16.8% of exposed adults (>17 years) were infected. All other studies used serologically identification (Table 2.9). As stated no information is available to qualify this. In this study we were unable to detect mild and/or sub-clinical infection due in part to the sampling interval and a large extent to the failure of the OF assay. Assuming some age-dependence of viral shedding it would imply that the study was less likely to have picked up infection in older children not because infection did not occur but because they were less easily detected. Hence it is likely that those infections indentified were probably measuring those cases shedding more virus – which is correlated to severity indicate more severe cases. However, these are the cases most likely to spread the infection – which is of epidemiological importance.

Majority of the infections identified were in pre-school children. Thus the hypothesis that older children especially school children are more often than not those who bring infection into the home [15, 16, 128, 144] in general did not hold true in this study as only a small

proportion (11%) of the infected siblings were attending school. 58% of the children in this study population were attending school. This increased chance of acquiring infection outside the home (from school contacts) would ideally be reflected in the number of index cases in this study who were in school. Indeed within the age class of potential school children (children 5 years of age and above), 67% of the index infections were from a child attending school (Table 5.6). These results are likely to be biased by the study design which favours detection of infection in younger children. Therefore, the observed reduced infection rates in older children in general may be a function of the clinical (symptomatic) case definition. Given results from previous studies the likelihood of missed sub-clinical infections in school children in this study cannot be ignored.

There were eight instances of more than one index case with two additional cases occurring within a day of first case, also considered to be index cases. Simultaneous cases within household members are generally rare as such cases require common exposure and transmission as well as similar likelihood of disease in the hosts or increased virulence in the pathogen. For such cases it is possible that one of the index cases was infected first. The data necessary to validate this (i.e. exact duration of shedding prior to clinical symptoms) is not available in this study. However, instructions given to the family may also play an important role in the timely identification of illnesses and in this study the importance of presenting to the OP clinic immediately symptoms were detected was emphasized. Hence we are re-assured that most samples would have been collected towards the start of clinical symptoms.

When dealing with a disease that is prevalent in the community as is the case with RSV during epidemics, it is not always possible to determine where an infection was acquired.

Accordingly identified cases of spread within the family may have been the result of independent exposure outside of the family and not from within the family. Nevertheless, documentation was made of several instances of secondary transmission in this study. Given the close contact occurring in families, it was expected that several family members would become infected once the virus was introduced as has been reported in nursery studies [119] and other family studies [15, 50]. Rather surprisingly in most instances an index case(s) resulted in only a single secondary case. Households provide a special setting as "outbreaks" in households imply transmission of infection under particular intense conditions of contact as well as the issue of extended physical closeness. The infectious dose thus would be higher increasing the probability of transmission taking place adding weight to the suggestion of missed infections.

The efficiency of transmission was investigated. The potential for spread is determined by probability of transmission taking place. This may in part be dependent on the level of susceptibility of exposed individuals, the contact patterns and to some extent the infectious dose. Thus, there are some contacts that are more likely to result in infection than others. Overall, several important patterns were observed i) infants in the family appeared to be infected earlier (significantly shorter interval between index and secondary case). As mentioned this may be the results of short incubation periods in younger child an aspect that remains unknown in the present study ii) SAR were higher for younger children as previously reported by Hall *et al* [15], and iii) older children seemed to be more effective at transmitting to younger children in the home than children of own age - older children preferentially infected younger children than similar aged children. Older children in this study seem to be less susceptible presumably due to previous history of infection (higher levels of antibody) but if infected are more effective at transmitting to younger children

who are more susceptible. Households of smaller size had significantly higher SAR presumably due to younger children. These households had a lower average age of children than did larger (6+ members) households (2-4 years vs. 5-7 years). As with all aspects of this section, these estimates are affected (biased) by possible age-dependence in detection of cases i.e. older cases less likely to be detected.

Of siblings <15 years, 23% had evidence of infection with RSV. 17% of the infections identified during follow-up were re-infections and these were in younger children (<4 years). The majority of these had re-infections identified in successive seasons. Interestingly, there were 4 instances where the re-infections occurred in the same epidemic implying that re-infection can occur within a short interval presumably an indication of short livid immunity discussed in Chapter 2 (Section 2.3.1) and published [110].

It is generally accepted that by three years of age most children have had a primary infection [12, 19, 25, 90, 111, 114]. Thus infections occurring in children/siblings 3 years of age or older were assumed to be re-infections. Using this definition, 49% of the infections in this study were in effect re-infections. Clearly the contribution of re-infections to transmission in this community - though in fact underestimated in this study - is not negligible. This is explored further in Chapter 7 in which we calculate the age-stratified incidence estimates. Several factors (discussed in Chapter 2) may contribute to the occurrence of re-infections [18, 19, 101] and include individual age, viral antigenic variation, host neutralizing antibody and of course the time between challenges. At the time of writing this thesis there was incomplete data available concerning antibody levels or data on the circulating strains during the study. Those samples that were genotyped showed that Group A was the dominant strain in epidemic 2 and co-dominant in epidemic 3

(unpublished data, DJ Nokes; Figure 2.4). 68% of re-infections observed were in birth cohort children between the ages of 1- 2 years. Age thus appears to affect clinically observed re-infection; 68 % of the re-infections were in children younger than 5 years compared to 30% in 5 to <10 year olds and only 2 % in those 10 years or older. This relationship of clinical versus sub-clinical infection and age can only be clarified using serological data.

Infected and non-infected households in this study did not differ in terms of family structure. They both exhibited similar patterns of pre-school children and school attendees, hence there was no apparent explanation for why certain households did not get infected, other than a purely stochastic process of introduction. Similar results have previously been documented [9, 15]. It may be that certain socio-economic characteristics were different for households in this study or the location in relation to the epidemic was different. However, there does not appear to be spatial clustering; since, both infected and uninfected households were from similar locations. Differences by socio-economic class were considered. Infected households were found to have a higher proportion of households classified as poor (72%) than did the non-infected households (41.2%), a difference that was found to be significant (Pearson χ^2 test; $p = 0.022$). A household's socioeconomic classification was also observed to influence the number of cases occurring after an introduction. Thus the reason for certain households being preferentially infected could be a predisposition effected by lower socio-economic status. Lower socioeconomic status has been associated with ARI morbidity; these households generally have poorer hygiene affecting the general health of members [259].

Lastly, it is important to note at this point that the results presented here do not give a complete picture of household transmission because of various factors. First, not all families studied had all members under routine clinical surveillance and even those that did were not under routine surveillance for the entire period of the study. Thus most of the data presented here is limited to transmission relating to children in the family. Second, the study protocol limited sample collection to symptomatic infection. An important adjunct of this was to be the results of surveillance for sub-clinical infection using oral-fluid (described in Chapter 4). Renewed efforts to optimization the oral-fluid ELISA are currently underway. However information concerning asymptomatic infection or infection not detected by our diagnostics method is to this extent still incomplete. Only 121 infections were identified in 373 participants followed up for two and half years with 56% of these being children ≤ 5 years and with only a single infection identified in an adult. Hence, it is likely that some RSV infections within this HH cohort were missed. Results presented in this chapter should be viewed in light of these limitations.

The primary aim of the birth cohort study, within which this study is nested, was to assess the effect of past history of infection on the risk of re-infection. For this it was desired to make a comprehensive record of RSV infection in each child and in particular to maintain follow up long enough to ensure a large proportion of children experienced more than one RSV infection (to allow the computation of protective efficacy - Chapter 1). Consequently, the recruitment strategy was designed for this purpose rather than for the estimation of incidence – for which continuous recruitment throughout the year would have been more useful in the estimation of finely age stratified rate of infection. One interesting consequence of the adopted intermittent recruitment (i.e. roughly 6 months of each year) arose in the second phase of the cohort as a result of the RSV epidemic starting 3 months

earlier than the preceding epidemic. The result was that only a small proportion of children experienced an RSV infection in their first 6 months of life. However, this did facilitate estimation of relative risk of RSV disease by age following primary infection. While it is acknowledged that intermittent recruitment creates problems in incidence estimation, the effect can be largely overcome by a process of weighting for age-specific observation time within/between epidemics as detailed in Chapter 7.

In spite of these deficiencies highlighted in the discussion, the data collected – which relate to clinically apparent infections only – have illuminated our understanding of various aspects of household transmission, and provided some thought provoking data regarding transmission patterns within the home. We identified substantial infection rates within households, documented infection introduction and defined interval between cases and SARs in relation to various factors. In the next chapter, analysis of these data also provides useful information on RSV transmission rates and risk of disease on primary and re-infection and the duration of shedding in relation to various factors.

Chapter Six

Duration of Viral Shedding

6.1 Introduction

The infectious period is a key characteristic determining the transmission of infectious pathogens (see schema Fig.5.1). In terms of RSV transmission dynamics it is important to quantify the infection / recovery process, and in particular to know how past infection modifies the infectiousness of individuals [60]. Because of the low average age at infection, the vast majority of the host population is in the category of 'previously infected'. The contribution of subsequent re-infections to the total rate of infection within a community will be highly dependent upon their duration of shedding. Equally, the occurrence of long term shedding of RSV may have significant implications to viral persistence, viral genetics and continued transmission [41]. Prolonged excretion of RSV enhances the possibility of transmission and makes such individuals potential sources of community spread of infection an important consideration in the control and prevention of RSV infection in infants. An assessment of this issue requires improved data on the prevalence and duration of viral shedding. The aim of this study is to estimate the duration of infectiousness. There are three durations commonly recognized in viral infections: from infection to the start of infectiousness (often called latent period in infectious disease epidemiology), from infection to onset of clinical symptoms (often called incubation period in clinical literature), and from onset of infectiousness to cessation of infectiousness (often called infectious period in epidemiological literature). All of these periods will have a distribution, i.e. will vary between individuals and infectious episodes. In particular, there will be a distributed period between infection with virus and the start of infectiousness for which we cannot have direct information in an observational field study. Further, infectiousness can only be assessed indirectly, that is, generally through detection of virus in nasal sections.

Consequently, we must assume that shedding is closely related to infectiousness (see discussion) and consequently, we are estimating the duration of viral shedding as a correlate for the infectious period distribution (IPD).

Estimation of the duration of shedding for perfectly observed data is straightforward. However, recognised in this study were two problems with the way the data were collected with a bearing on the estimation of shedding duration. First, recruitment into the study was based on a positive sample so that the date of onset of shedding is unknown, i.e. there is left censoring due to the recruitment protocol. The information we have defines an interval during which shedding started, i.e. the date of last known negative and the date of first positive. Second, the exact date of cessation of shedding is not known. The data defines the interval in which cessation of shedding occurred as samples were obtained at three day intervals. If cessation of shedding is defined as failure, then this is termed interval censoring. This was the case for all observations in this study except for one individual who died before end of shedding, i.e. was right censored. This individual's data was excluded from the analysis. As the intervals between sampling were relatively short, the interval censoring is not critical but to be correct should be accounted for in the analysis. However, some of the left censored intervals are large and their exclusion could substantially alter the analysis outcome and give rise to spurious results. It is acknowledged that there is a probable third problem that is the single negative indicating failure rather than confirmed by two negatives (see discussion).

A review of previous studies on the duration of shedding has been given in Chapter 2. Typically, RSV has a short infectious period of approximately 6 days. However, the distribution of shedding duration is skewed with a small proportion of individuals shedding

for relatively longer periods. It is therefore interesting that the disease is able to persist even with such a short duration of infection. The hypothesis advanced to explain the dynamics of shedding is that infected individuals fall into two categories i) children suffering from a primary infection and who are potentially longer shedders but form a small proportion of the population and ii) those with a re-infection (majority of the population) but who are thought to shed virus for a shorter duration i.e. each re-infected person less important than a primary case (Fig.5.1). It is not clear exactly which of the two groups is the major contributor and thus potentially more important in RSV transmission dynamics. Re-infection occurs frequently (Chapter 2), and is probably very important in overall transmission. If indeed re-infections have a shorter duration of shedding than primary infection, then the magnitude of this shedding reduction would greatly affect the total contribution of these re-infections to the pool of infection. An attempt was therefore made to quantify this.

6.2 Chapter aims

In contrast to much of the information available regarding shedding of RSV (from hospitalised patients), in this Chapter we give a report of the duration of shedding of RSV from infected individuals in the community, in relation to past infection history, i.e. primary or re-infection, by age, sex and by the severity of infection.

6.3 Methods

Study population and samples

A study to determine the duration of shedding of RSV after infection was nested within the surveillance framework of the main birth cohort project and the household cohort study (described in Chapter 3). The study observation period reported for the duration of

shedding study is that from two epidemics between January 2004 and June 2004 and November 2004 and February 2005. Nasal washings were collected routinely from participants experiencing episodes of acute (rapid onset) respiratory illness which were screened for RSV by commercial direct immunofluorescent antibody test as described in Chapter 2. During 2004 and 2005 any child in the cohort identified (through either home visit surveillance or passive referral) as RSV positive was asked to enrol in the shedding study. Following the identification of RSV infection (day 1 or t_0), a further nasal washing was obtained as soon as possible, and thereafter every 3 days up to day 14. Individuals remaining RSV positive on their 4th sample were followed-up and sampled after a further 7 and 14 days, and thereafter every 2 weeks up to 16 weeks (Figure 6.1). Follow up was discontinued following a single sample testing negative.

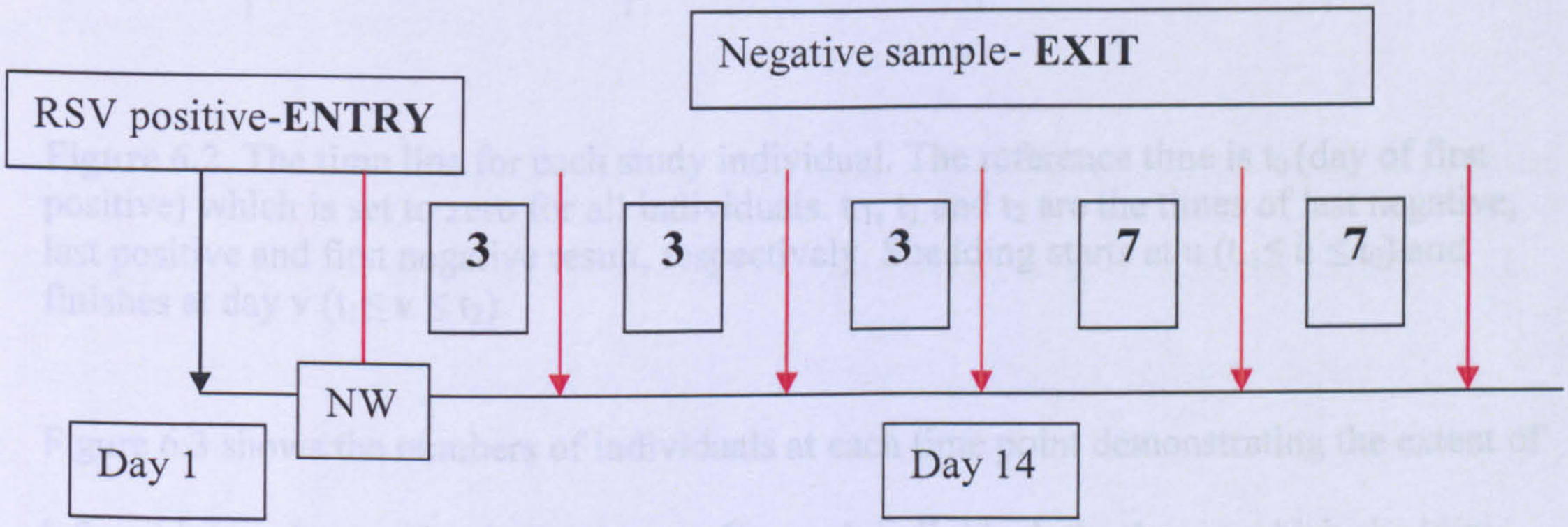


Figure 6.1. Illustration of study design. Red arrows show sample collection.

Data Analysis

Data was analyzed using Excel (Microsoft Office) program and Matlab (Release 14, www.matlab.com). Relevant observations for each individual, i, are sample dates of most recent prior virus negative, first virus positive, last virus positive and first virus negative: $t_{-1,i}$ t_0, $t_{1,i}$ and $t_{2,i}$ respectively (Figure 6.2). The prior negative sample refers to the date of the

most recent sample collected which was antigen negative before the positive sample (first virus positive) that initiated shedding follow up in the individual. All data are in days, and we assume that all samples were taken at noon of each day. The time of first positive sample, t_0, is set to zero and so all observed times are relative to that time. For each individual, the start day of shedding, u_i, and the day of cessation of shedding, v_i, are defined by: $t_{-1,i} \leq u_i \leq t_0$ and $t_{1,i} \leq v_i \leq t_{2,i}$ respectively. Where the first prior observation was positive and not negative (occurring in 2 instances), $t_{-1,i}$ is set to -200d.

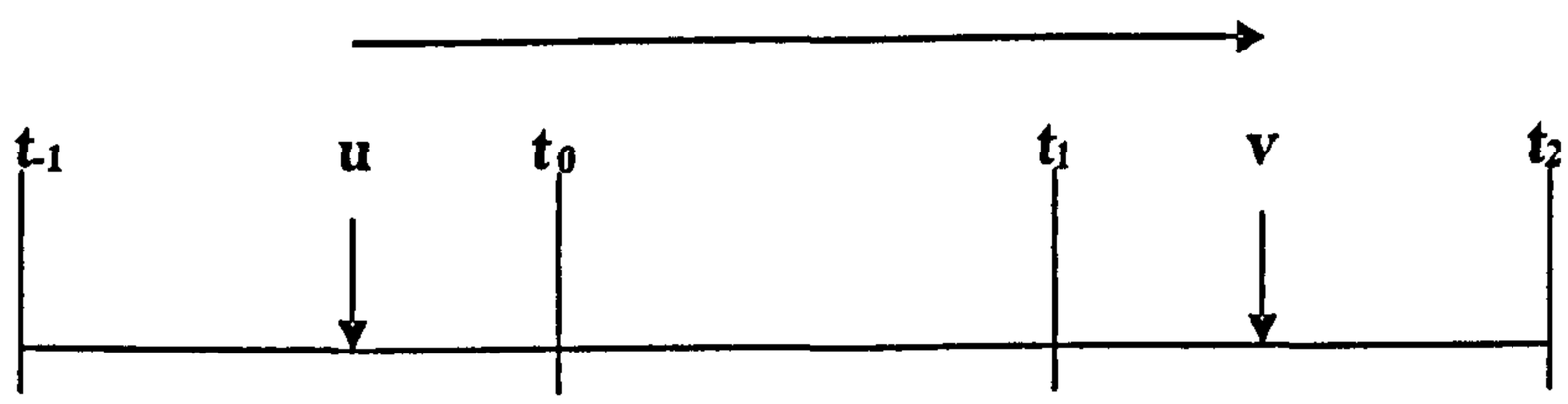


Figure 6.2. The time line for each study individual. The reference time is t_0 (day of first positive) which is set to zero for all individuals. t_{-1}, t_1 and t_2 are the times of last negative, last positive and first negative result, respectively. Shedding starts at u ($t_{-1} \leq u \leq t_0$) and finishes at day v ($t_1 \leq v \leq t_2$).

Figure 6.3 shows the numbers of individuals at each time point demonstrating the extent of left and interval censoring in the dataset. For each individual, the day on which shedding was first detected is taken as the reference time (t_0= day 0). The left hand distribution shows the number of individuals at each day prior to detection of shedding who might have been shedding on that day, i.e. between the last negative result and the first positive result. The right hand distributions show: i) the number of individuals at that time point known to still be shedding virus (a positive nasal wash after first positive (t_0); $t \leq t_1$- dark grey bars) ii) the number of people known to have stopped shedding (confirmed negative nasal wash after t_0; $t \leq t_2$ - light grey bars), iii) the number whose status is unknown (nasal washing not yet collected- illustrating the interval between sampling- yellow bars).

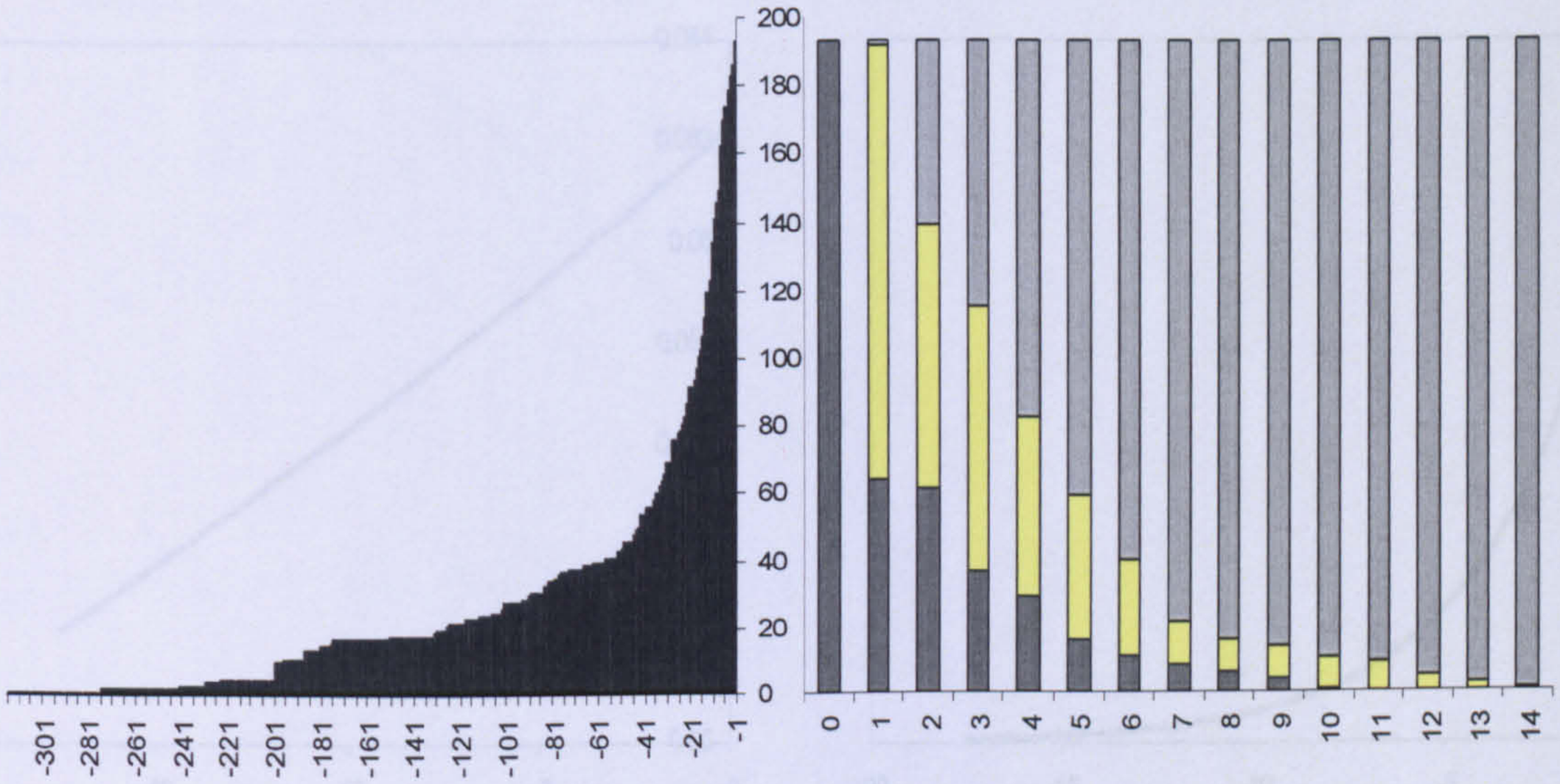


Figure 6.3. The number of individuals at each time point. For each day is shown the number known to be shedding (dark grey), known to have stopped shedding (light grey), of uncertain classification due to interval censoring (interval between positive and negative samples) (yellow) and at risk of having started shedding, i.e. <day 0 (black bars). The horizontal axis is time in days, with t=0 as the reference for each individual. (left hand side of Figure has been modified).

The two types of unobserved times (infection onset and cessation) need to be accounted for in the estimation. We assume that the rate of commencement of shedding is a constant, λ, so that the distribution of start times follows a truncated exponential distribution:

$$f(u_i \mid \lambda) = \frac{e^{\lambda u_i}}{1 - e^{\lambda t_{-1,i}}} \quad \lambda > 0$$

$$= \frac{1}{-t_{-1,i}} \quad \lambda = 0$$

noting that with our definition of time $t_{-1,i}$ and u_i are negative. For high values of λ the probability of infection becomes increasingly concentrated towards t_0 (Figure 6.4).

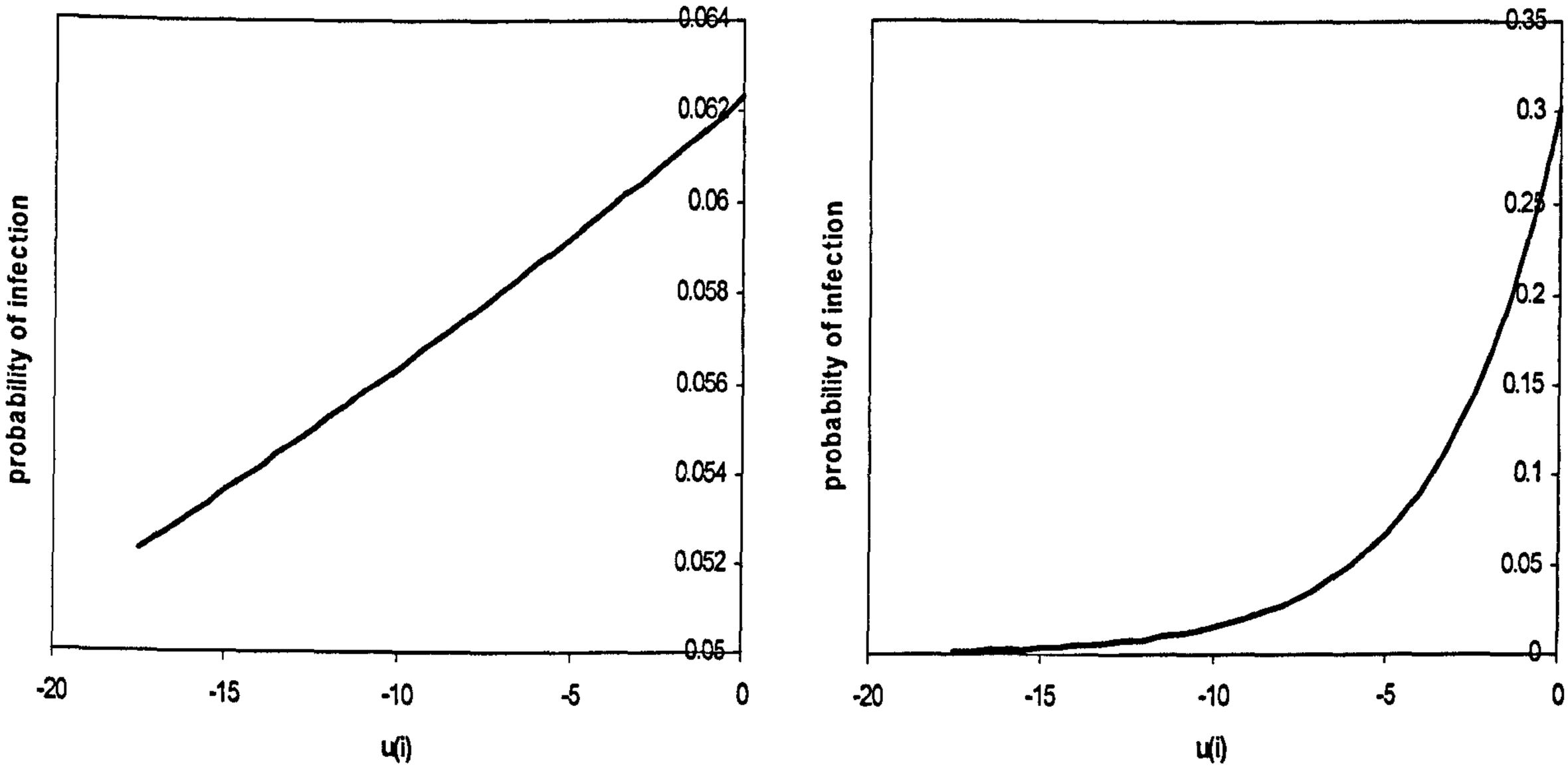


Figure 6.4. Probability of infection in individual i at different times, u (days) assuming a truncated negative exponential distribution and infection occurs before day 0. a) λ =0.01; b) λ =0.3

The Weibull probability distribution function was used to estimate the probability density function of shedding duration. This distribution is sufficiently flexible to capture the important features of the observed shedding times, and relatively easy to handle computationally. The probability density function is given as

$$P(t \mid \alpha, \beta) = \alpha\beta t^{\beta-1} e^{-\alpha t^{\beta}},$$

and the survival function as

$$S(t \mid \alpha, \beta) = e^{-\alpha t^{\beta}}$$

where α and β are two unknown parameters to be estimated. Figure 6.5 shows a particular Weibull distribution.

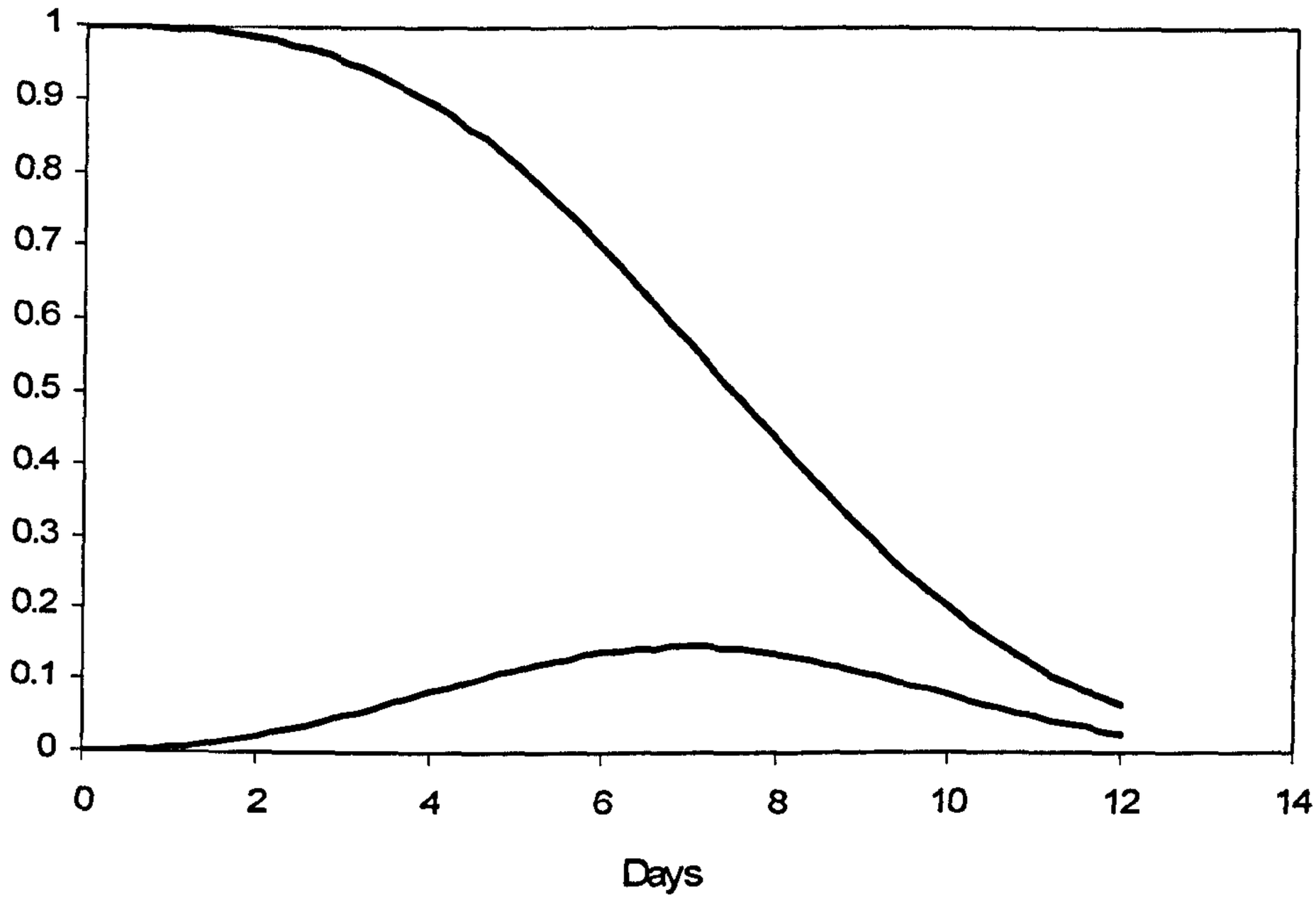


Figure 6.5. The Weibull probability distribution as the survivor function (blue) and the probability density function (black) ($\beta = 3$ and $\alpha = 0.002$). As the shape parameter (β) increases the survivor function becomes increasingly like a step function.

The mean (m) and variance (σ^2) of the distribution are given as

$$m = \left(\frac{1}{\alpha}\right)^{1/\beta} \Gamma\left(1 + 1/\beta\right)$$

$$\sigma^2 = \left(\frac{1}{\alpha}\right)^{1/\beta} \left[\Gamma\left(1 + 2/\beta\right) - \Gamma\left(1 + 1/\beta\right)^2\right]$$

where Γ is the gamma function.

If the start of shedding (u_i) was perfectly observed, then the likelihood of observing shedding at $t_{1,i}$ and no shedding at $t_{2,i}$ is:

$$L_i(u_i, \alpha, \beta) = \left\{S(-u_i + t_{1,i}) - S(-u_i + t_{2,i})\right\}$$

We compute the full log-likelihood by integrating over all possible values of u_i, and summing over all individuals:

$$L(\alpha,\beta\,|\,\lambda)=\sum_{i}\ln\int_{t_{-1,i}}^{t_0} f(u\,|\,\lambda)\{S(-u+t_{1,i})-S(-u+t_{2,i})\}du$$

This likelihood was computed numerically in MatLab (Release 14, www.matlab.com) by Prof. Graham Medley.

The relationship between duration of shedding and age, sex, severity of infection and history of infection was evaluated (Table 6.1 and 6.2). All factors considered had two categories. We compared models using the likelihood ratio test (LRT) based on the likelihood ratio statistic (LRS) = $-2(L_0-L_1)$. Table 6.1 shows the numbers of individuals in each category tested.

6.4 Results

A total of 193 RSV positive children were enrolled into the shedding study, 160 were birth cohort infants and 33 siblings from the household study. 192 failures (end of shedding) were observed, 1 child died before completing the study and was excluded from the analysis. The children were between the ages of 2 and 164 months (i.e. 0.2 - 13.7 years); 9.5% were less than 1 year while 71.9% were 2 years of age or less. The median age of children in the study was 30 months, and 45.4% were male. Of the 193 RSV infections 165 were classified as an upper respiratory tract infection (URTI), 20 as having mild LRTI and 7 with severe LRTI, none with very severe LRTI.

For 120 of the children, the day on which they were found to be positive was equal to the day of their last positive test ($t_0 = t_1$), indicating that the majority of children only shed virus for a maximum of 3 days (Figure 6.3). To get some insight into the data, Kaplan

Meier curves of survival assuming infection at t_0 with failure as first negative were plotted.

Only plots with highest significance are shown.

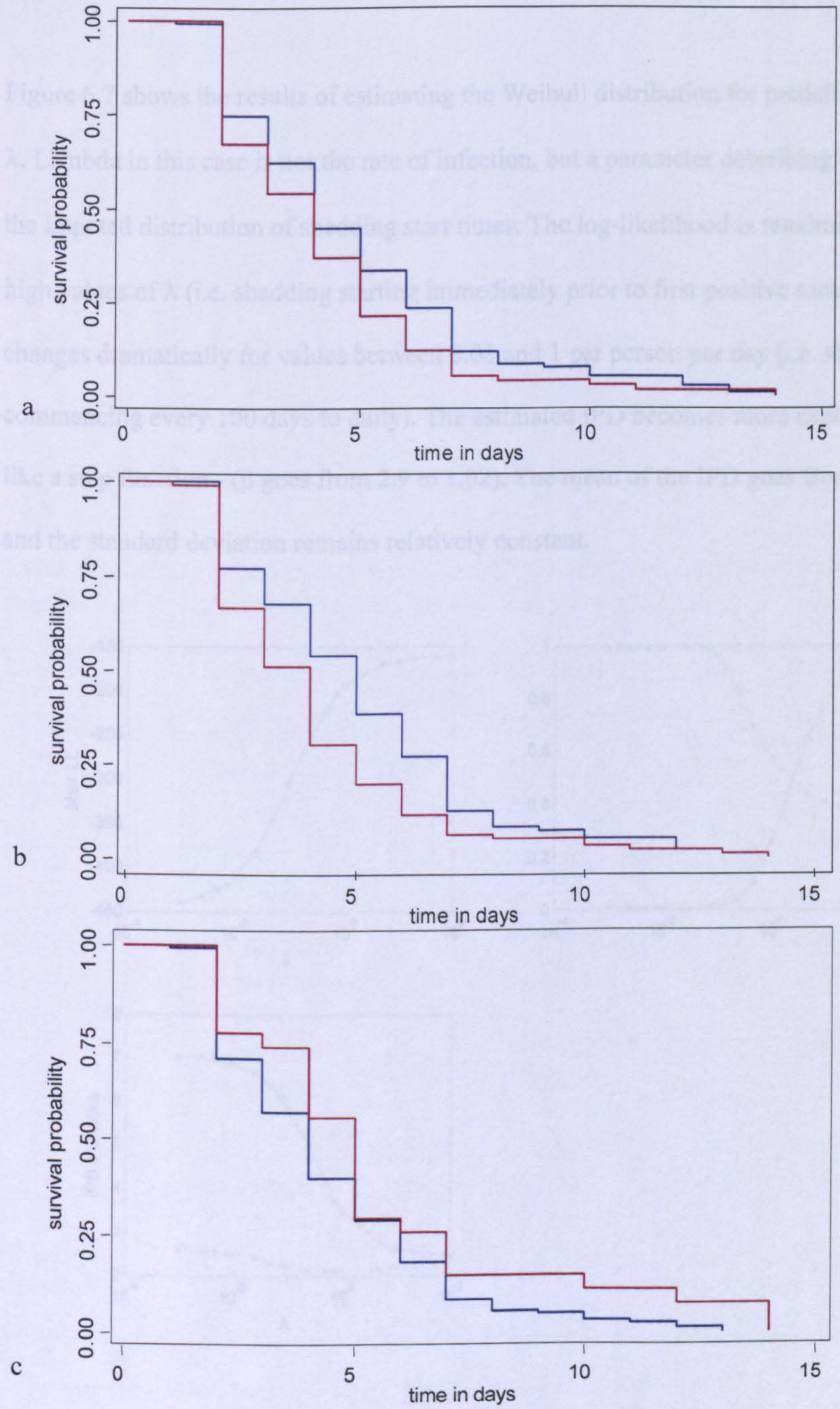


Figure 6.6. Kaplan Meier survival estimates for covariates with highest significance a) Age class (blue: 0-23months; maroon: 24+ months) b) Infection history (blue: never infected; maroon: infected- significantly different; log rank test $p<0.05$) and c) severity (blue: URTI; maroon: LRTI).

Figure 6.7 shows the results of estimating the Weibull distribution for predefined values of λ. Lambda in this case is not the rate of infection, but a parameter describing the shape of the imputed distribution of shedding start times. The log-likelihood is maximised for very high values of λ (i.e. shedding starting immediately prior to first positive sample) and changes dramatically for values between 0.01 and 1 per person per day (i.e. shedding commencing every 100 days to daily). The estimated IPD becomes more exponential – less like a step function - (β goes from 2.9 to 1.02). The mean of the IPD goes from 7d to 2.5d and the standard deviation remains relatively constant.

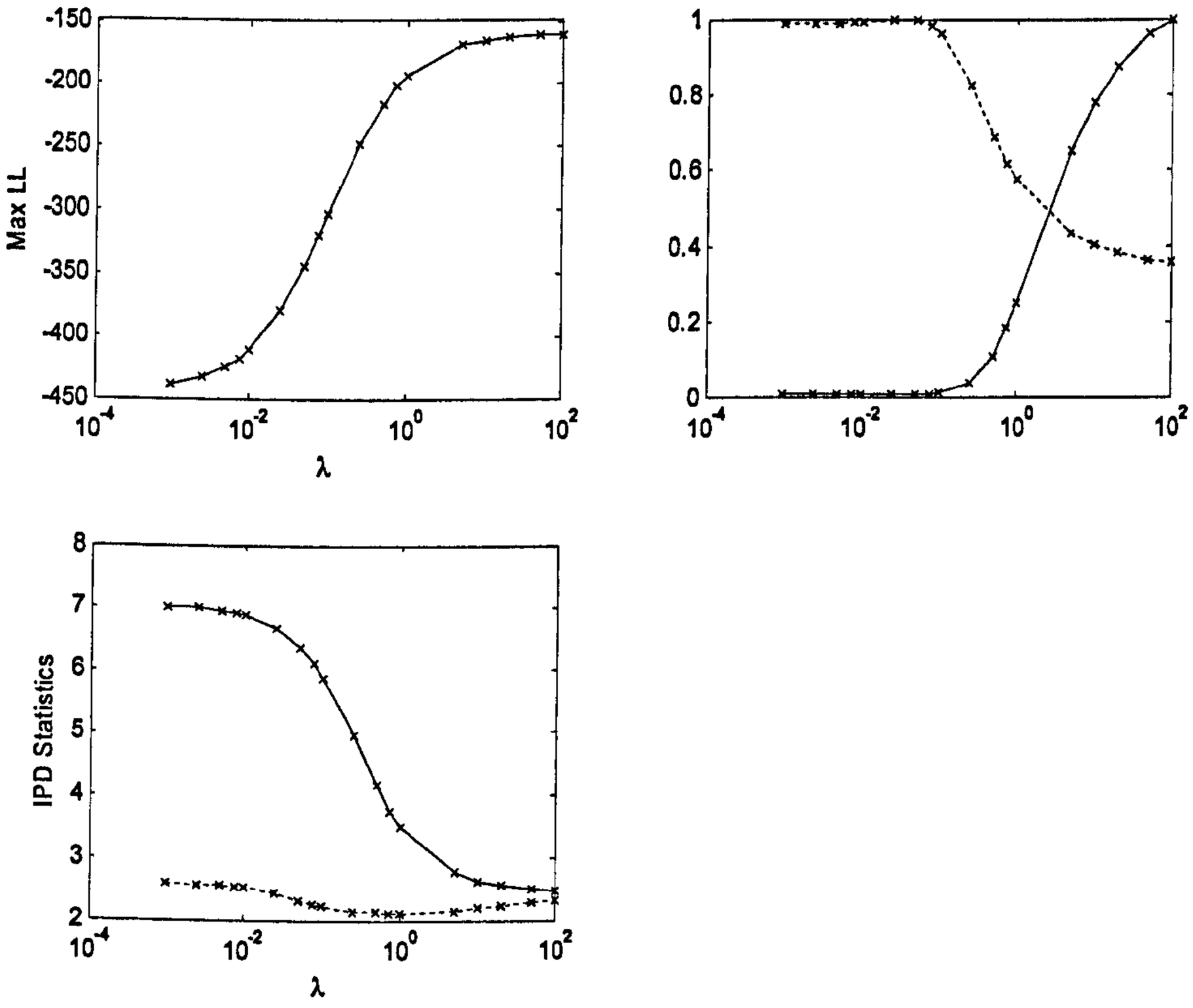


Figure 6.7. Variation in results with **model parameter (λ)** (λ). a) the maximum log-likelihood (top left); b) the parameters of the Weibull IPD function relative to their maximum value (solid line: α; dashed line: β); c) the mean (solid line) and standard deviation (dashed line) of the IDP (bottom left).

The two tables below (Table 6.1 – best fit model and Table 6.2) show the results for the 6 binary variables that were investigated using two values of lambda (λ=10 and λ=0.01). None of the variables examined were found to significantly affect the duration of shedding. The estimation procedure was confirmed by adding a dummy variable to divide the data on the basis of the observed duration of shedding, and was found to be working. Although no formal power calculations have been performed, if there were a consistent, large effect of any of these variables it would be detected by this approach.

Table 6.1. Infectiousness period distribution (IPD) estimated using survival analysis using data from 192 infected children from the best fitting model (λ=10).

Factors	**Categories**	**Numbers N=192**	**Mean IPD**	**beta**	**LL***	**LRS[$]**
History	never infected	96	2.904	1.337	-163.469	2.964
	infected	96	2.362	1.106		
Sex	male	88	2.495	1.166	-164.658	0.586
	female	104	2.760	1.245		
Age class	0-11 months	20	2.534	1.302	-164.859	0.184
	12+ months	172	2.656	1.200		
	0-17 months	55	2.652	1.277	-164.839	0.224
	18+ months	137	2.634	1.179		
	0-23 months	121	2.817	1.248	-164.103	1.696
	24+ months	71	2.338	1.153		
Severity of infection	URTI	165	2.502	1.202	-163.293	3.316
	LRTI	27	3.478	1.325		

LL- Log-likelihood
*The Log-likelihood for the full (saturated) model is -164.951
[$]1 df at 5% significance the chi-squared value is 3.84. Parameter values with LL above the line at 3.84 are not statistically signifcant

Table 6.2. Infectiousness period distribution estimated using survival analysis using data from 192 infected children with λ=0.01.

Factors	Categories	Numbers N=192	Mean IPD	beta	LL*	LRS
History	never infected	96	6.996	3.479	-407.499	1.458
	infected	96	6.69	2.636		
Sex	male	88	6.634	2.768	-407.838	0.780
	female	104	7.078	3.209		
Age class	0-11months	20	5.571	2.848	-406.965	2.526
	12+ months	172	7.076	3.075		
	0-17 months	55	6.647	3.143	-408.071	0.314
	18+ months	137	6.984	2.899		
	0-23 months	121	6.916	2.908	-408.185	0.086
	24+ months	71	6.816	3.099		
Severity of Infection	URTI	165	6.659	2.986	-407.162	2.132
	LRTI	27	8.11	3.199		

LL- Log-likelihood

*The Log-likelihood for the full (saturated) model is -408.228

6.5 Discussion

In contrast to much of the information available regarding shedding of RSV, in this thesis we give a report of the duration of shedding of RSV from infected community individuals, in relation to past infection history, i.e. primary or re-infection, by age, sex and by the severity of infection.

Sensitivity of detection of shedding (infectiousness) is likely to be determined by intensity of shedding, which is expected to increase then decrease as the infection proceeds within the individual. It may be assumed at the early stages of virus replication, viral load in secretions will be low, and it has been shown [63] that as the days post (samples collected soon after admission) infection increase subsequent specimens tend to have lower RSV

viral load. Consequently sensitivity of antigen detection will change during the infectious period: individuals first assessed during the early or late stages of infectiousness will be less likely to be included (a potential bias in recruitment) into the study. The second bias is that of identification of failures which was assumed to arise after a single negative result. Given the nature of error in biological measurement – particularly at low level shedding towards the end of the infection, it is conceivable that children testing negative one day might subsequently be positive on the next. Ideally we would have used the criterion of two sequential negatives as indicative of a failure. However, given this argument (that we are dealing with the terminal part of the infection with low viral load) it is most likely that the probability of a false negative is low. Indeed in a sub-sample (102) we did change the criterion to double negative for a failure, and this gave rise to a risk of false negativity of 4.9%. Future analysis planned will account for this small bias in the estimation procedure. These are additional to those which we have attempted to take account of in the analysis i.e. uncertainty of the exact times of onset and cessation of infection, u and v. Data on intensity of positives was not available, and hasn't been included.

For a fuller analysis we would need to take account of additional information concerning the force of infection for RSV. Since the sampling interval was decreased during periods of high RSV activity (from 4 to 1 week) the enrolment is potentially biased towards an increase in estimated length of time elapsed from the point of infection to recruitment. This would arise for those cases occurring early in the epidemic. Also, if sub-groups of the population e.g. older children have shorter IPD, then they are less likely to be included in the study. This aspect will, also, potentially bias the analysis if it is not included. In particular, inclusion of calendar time and external information on RSV activity (i.e. from IP surveillance) will enable potential biases to be included in the analysis. In particular, the

duration of IPD is directly related to the probability of inclusion in the study, so that the analysis is biased against shorter IPD.

The estimated duration of shedding (Fig 6.7) is a few days (2.5-7days). This is in accordance with previous results [15, 57-59]. The best fitting model was apparently for this duration to be exponentially distributed with mean approximately 3 days, There is little information in the left censored times to indicate what should be imputed. Given the invasive nature of sampling, there is little that can be done to improve the sampling intervals related to recruitment. The analysis shows that the best models are with a zero imputation (i.e. individuals start shedding when they are recruited). This is unlikely, but information to qualify this for this study is currently unavailable.

Contrary to expectation [59, 62], the tested co-variates provided no statistically significant evidence of an impact on duration of shedding. Transmission dynamic modelling [60] has demonstrated that shortened duration of shedding during second and subsequent infections with RSV can account for many of the population level dynamic patterns. However, it might be that there is an inherent bias in the data collection against detecting shorter shedding periods. The KM plots in Figure 6.6 give insight into the data. The patterns of shedding in general appear consistent with assumptions about increased shedding duration in primary versus reinfected children and in those more severely infected. This ought to be further investigated.

Other methodological issues that might have affected the estimated duration include the assumption that all children not followed up from birth but who were over the age of three have had a previous infection. This assumption is warranted because by the age of three

years the great majority of children will have experienced a primary infection [12, 19, 25, 90, 111, 114]. Only studies that take samples irrespective of symptoms potentially give a precise estimate of the duration of shedding [59]. In this study by Frank *et al.* [59], any samples taken 8 days pre-illness were used to compute duration of shedding. Encouragingly, most isolates of RSV were obtained during the first seven days after the onset of illness (i.e. very few pre-symptomatic samples are positive). Thus, the fact that samples are taken only when symptoms are present in this study is not a big concern. An improvement of the present study design would be to collect daily samples pre and post-infection something that we believe would be difficult to justify and even implement (in terms of participant refusal). A possible compromise would be to take samples at closer intervals. Also, more sensitive detection techniques such as PCR might be warranted – however, it would be less certain whether identification by PCR relates to viable virus compared with the IFAT diagnostic method used here (i.e. observation of infected cells).

In conclusion, the duration of shedding of respiratory viruses in naturally infected persons with non-severe infection is important for understanding the spread of infection in a population and will contribute to the development of transmission dynamic models to investigate the impact of immunization. This chapter provides such estimates specifically taking account history of infection. Estimates of the duration of shedding remain conservative even though we have attempted to account for the left and interval censoring. There was no convincing evidence for an association between suspected risk factors and increased period of shedding.

Chapter Seven

Age-stratified incidence of infection and disease severity

7.1 Introduction

The epidemiology of RSV is mainly about infection and re-infection, and how this relates to disease. Community and hospital-based studies have reported the importance of RSV as a cause of LRTI and severe LRTI [4-9, 11, 205]. However, there are few disease burden studies from developing countries [38]. RSV-LRTD (pneumonia and bronchiolitis) is almost entirely confined to children less than 3 years of age [12, 16, 19, 35, 90, 94, 101, 260]. The risk of severe RSV infection is correlated with a number of factors , in particular, primary infection in early life, in children with underlying predisposing risk factors such as prematurity (presumably because of absence of maternal antibody), and with cardio-pulmonary disease. Some environmental factors such as exposure to smoke or socio-economic classification (explored in Chapter 8) have also been reported to increase the risk of RSV disease.

Though largely overlooked, because of the importance of infection in children, RSV causes repeated infection throughout life. One of the main characteristics of transmission is the occurrence of re-infection. This occurrence of infection in older age groups discussed in section 2.6 (Chapter 2) has important implications for community transmission but has as yet not been comprehensively investigated, and particularly so in developing countries where no study has explored the contribution of re-infection in older age groups and thus forms the basis of this Chapter.

7.2 Chapter Aims

This chapter aims to quantify the occurrence of ARI and RSV associated ARI by age, sex and child class (pre-school or school child) in the household cohort. The risk of disease following RSV infection is also investigated.

7.3 Methods

Study design

A total of 73 families of birth cohort children within the DSS were recruited between February and May 2003. Sampling was purposive rather than random. A sub-sample of households of birth cohort infants with at least one sibling were selected. Within these families, all children up to and including those aged 14 years were monitored for ARI for a maximum of 26 months. From the end of January 2004 consenting adults were also included. Monitoring involved active and passive surveillance as described in chapter 3. Nasal specimens were tested for RSV antigen by Direct Immunofluorescence. There were 25 households lost to follow up during the study, of which 8 were replaced before experiencing an epidemic (in 2004). Thus in total 81 (73+8) households were recruited.

Data and definitions

Data collected are the cases of ARI and RSV positive ARI (RSV-ARI) over the study period. At each active visit a record was made of absence from the District since the last visit. The presence of any respiratory symptoms was considered an illness episode (refer to Chapter 3). Due to the diverse aetiology of ARI, new ARIs were delineated if symptoms were present one week after the previous illness episode. Observation time was time at risk, that is, time present within the DSS during which episodes were potentially observable. Non-observational time was excluded from time at risk. Only children for whom a single

visit following recruitment arose were included as study participants. For a specified period following RSV infection (14 days) an individual was assumed not to be at risk of another infection hence this should be viewed as non-observational period and was excluded. The one week post infection was not excluded from observation time for ARI estimates. The last visit was treated as the exit date leading to the exclusion of observation time between the last visit and the documented exit time.

The severity of respiratory disease was assigned on the basis of WHO criteria as described previously [5]. Severe ARI (LRTI) is defined as a history of acute cough or difficulty in breathing and ≥ 1 of the following: (1) fast breathing for age ($\geq$60 breaths/min if <2 months old or $\geq$50 breaths/min if 2–11 months old or $\geq$40 breaths/min if 12+ months), (2) indrawing, or (3) low oxygen saturation (<90%) by pulse oxymetry or inability to feed (in practice defined as prostration or unconsciousness), and if accompanied by a clinical diagnosis of LRTI or bronchiolitis (this caveat applies only to the diagnosis of very severe LRTI). Severe LRTI was diagnosed when a child had LRTI meeting criteria 2 and/or 3 above.

Children were categorized as follows: a) attending school b) a pre-school child with a sibling(s) who attends school or c) a pre-school child with no sibling(s) who attend school. Individuals were combined into three age classes - infants, young non-school children (1-5 years), older school age children (6-14 years), and adults to calculate adjusted independent estimates. The definitions of an epidemic and inter-epidemic periods have been described previously in Chapter 3.

Analysis

Analysis was carried out using STATA 8. Incidences were estimated as the number of cases of ARI and of RSV-ARI per 1000 person years of observation. Poisson regression commands were used. Incidence rates and 95% confidence intervals for ARI and RSV-ARI and RSV-LRTI were calculated by age, sex and class of children in the household. To account for repeated episodes of ARI within individuals (significant within individual clustering), Poisson regression accounting for clustering using a random intercepts model was used. The same was not done for RSV infection since there very few reinfections. 95% confidence intervals for the incidence estimates were calculated using the quadratic approximation to the Poisson log-likelihood for the log-rate parameter. Adjusted (for age, sex and/or child class) incidence rate ratios (IRRs) (estimated by use of Poisson regression) between groups (95% CI) were calculated. Wald test (i.e. log (rate/SE)) of the null hypothesis that a particular rate ratio is 1 was used to assess for significance within groups and is based on the quadratic approximation to the log-likelihood ratio at the maximum. The Likelihood Ratio Test (LRT) was used to assess the presence of interactions between variables and to test the significance of a variable while controlling for others.

For the analysis of incidence of RSV infection and disease the follow-up time was split according to epidemic and non-epidemic periods. The nature of recruitment (restricted to 6 months and not all year round, a factor of the study design – see section 3.3.1 Chapter 3) resulted in a bias in observation time during epidemic and non-epidemic time periods in the different age classes. As a result time at risk of RSV infection was uneven between age strata. To account for this, the observation time was weighted by a ratio calculated using the actual and expected child years of observation (cyo) within and between epidemics, as follows:

Let the actual period of observation (days), *Y (i)*, for age class *i=1, m* be the summation of observation in age class *i*, for epidemic, *k=e,* and non-epidemic, *k=n,* periods

$$Y(i) = Y_n(i) + Y_e(i) = \sum_k Y_k(i)$$

The expected observation time in days, $E_k(i)$ for age class, *i*, is calculated for non-epidemic, *k=n* and epidemic, *k=e* periods as

$$E_k(i) = P_k . Y(i)$$

where P_e is the expected proportion of a year with epidemic and P_n is expected proportion of year without epidemic.

Hence, weights by age class *i* applied to data from within (*e*) and between (*n*) epidemics, $W_k(i)$, are the ratio of expected to observed child years of observation for each age class, *i*,

$$W_k(i) = E_k(i) \Big/ Y_k(i)$$

The expected proportions of the year within and without epidemic (P_k) were estimated to be 32% (16.3/52) and 68% (35.3/52) respectively. Using this weighting incidence was calculated with respect to age class and time as detailed in Appendix O (p 351).

7.4 Results

There were 73 families with 373 members who had more than one visit who are considered in this analysis. This included 314 children under the age of 17 years (<15 years at the start of the study) with a median age of 4 years. Most of the adults were females (50 of the 59). Participants numbering 87 (78 from whole HHs) dropped out from 25 families before the end of the study; 51% refused (i.e. withdrew consent), 44% moved out of the study area while the remaining 5% either died or were requested to exit the study. Those requested were as a result of a split in one polygamous household. 8 households were replaced. The

age and sex distribution of study participants at recruitment is given in Table 7.1. There were 137 children attending school, 4 of whom were classified as adults (i.e. > 14 years at recruitment), 133 pre-school children with one or more siblings attending school and 48 pre-school children with no sibling(s) in school.

Table 7.1. Age and sex distribution of study participants at recruitment

	Sex		
Age (years)	**Female**	**Male**	**Total**
<1	50	45	95
1	7	6	13
2	20	20	40
3	12	12	24
4	7	12	19
5	11	9	20
6	10	10	20
7	9	7	16
8	3	3	6
9	4	9	13
10	10	10	20
11-14	14	14	28
15+ (adults)	50	9	59
Total	**207**	**166**	**373**

7.4.1. Incidence of all respiratory illnesses

A total of 2,566 ARIs were recorded from a total observation time of 537.3 years with an average of 477 respiratory illnesses per 1000 person-years (95% CI, 459 -499 illnesses/ 1000 person-years). There were 203 cases of LRTI identified in the household cohort with an incidence of 378 per 1000 person-year (CI 330-430).

i) Age-specific incidence rates (IR)

There was a general decline in all-cause respiratory illness rates with age (Table 7.2). There was a general reduction in the rate of illness from 2 years of age except for a slight increase observed in children aged 3 years. Adults had fewer attacks compared to children, with higher rates observed in women (1.59 (95%CI 0.98, 2.60) vs. 0.79 (0.30, 2.06)). The estimate of the incidence of LRTI in < 5 was 679 cases/1000 cyo (95% CI, 550-839), that in < 15 was 410 cases/1000 cyo (95% CI, 327-514).

Table 7.2 Incidence of respiratory illness according to age

Age (years)	# of illnesses	Person years of observation	ARI per person-year (95% CI)*	ARI per person-year (95% CI)* adjusted for clustering of episodes
<1	596	75.7	7.87 (7.26, 8.53)	7.62 (6.71, 8.66)
1	662	67.6	9.79 (9.07, 10.57)	9.41(7.39, 11.99)
2	202	38.6	5.68 (4.94, 6.51)	5.29(3.80, 7.37)
3	328	47.3	6.92 (6.22, 7.72)	6.60 (4.77, 9.13)
4	186	40.1	4.64 (4.02, 5.36)	4.74(3.34, 6.72)
5	106	29.8	3.56 (2.94, 4.30)	3.44 (2.33, 5.09)
6	103	31.2	3.30 (2.72, 4.00)	3.46 (2.32, 5.15)
7	79	30.1	2.62 (2.10, 3.27)	2.74 (1.80, 4.18)
8	59	22.2	2.66 (2.06, 3.44)	2.56 (1.62, 4.04))
9	26	14.9	1.74 (1.19, 2.56)	1.78 (1.00, 3.17)
10	31	18.8	1.65 (1.16, 2.34)	1.72 (0.99, 2.96)
11	21	22.2	0.95 (0.62, 1.45)	0.97 (0.53, 1.78)
12	28	16.5	1.69 (1.17, 2.45)	1.61 (0.92, 2.83)
13	35	15.2	2.31 (1.66, 3.21)	2.36 (1.37, 4.07)
14	17	9.1	1.87 (1.16, 3.00)	2.10 (1.05, 4.20)
adults	87	60.9	1.43 (1.16, 1.76)	1.45 (0.97, 2.19)
Total	**2566**	**537.3**	**4.77 (4.58, 4.96)**	**4.53 (4.11, 5.00)**

*95% confidence intervals for ARI incidence are calculated using the quadratic approximation to the Poisson Log likelihood for the log-rate parameter.

ii) Incidence of respiratory illness by age and gender

Infection rates were compared by age and gender. Among children 1 year old or younger, the incidence rates were higher in male children than in females (Figure 7.1). IRR with age class and sex were calculated. There was a significant age effect in some age classes and not in others suggestive of an interaction. The results show that the interaction between sex and age is not significant (lrtest, p=0.122). School children (6-14 years) and adults had significantly lower rates after adjusting for sex; 0.41 and 0.24 respectively. The rate in pre-school children (1-5 years) was the same as that observed in infants (IRR 1.06, p=0.246). Although sex does not appear to be significantly associated with respiratory infection rate, among adults the rate was twice as high in females (IRR =0.49) but this was not significant.

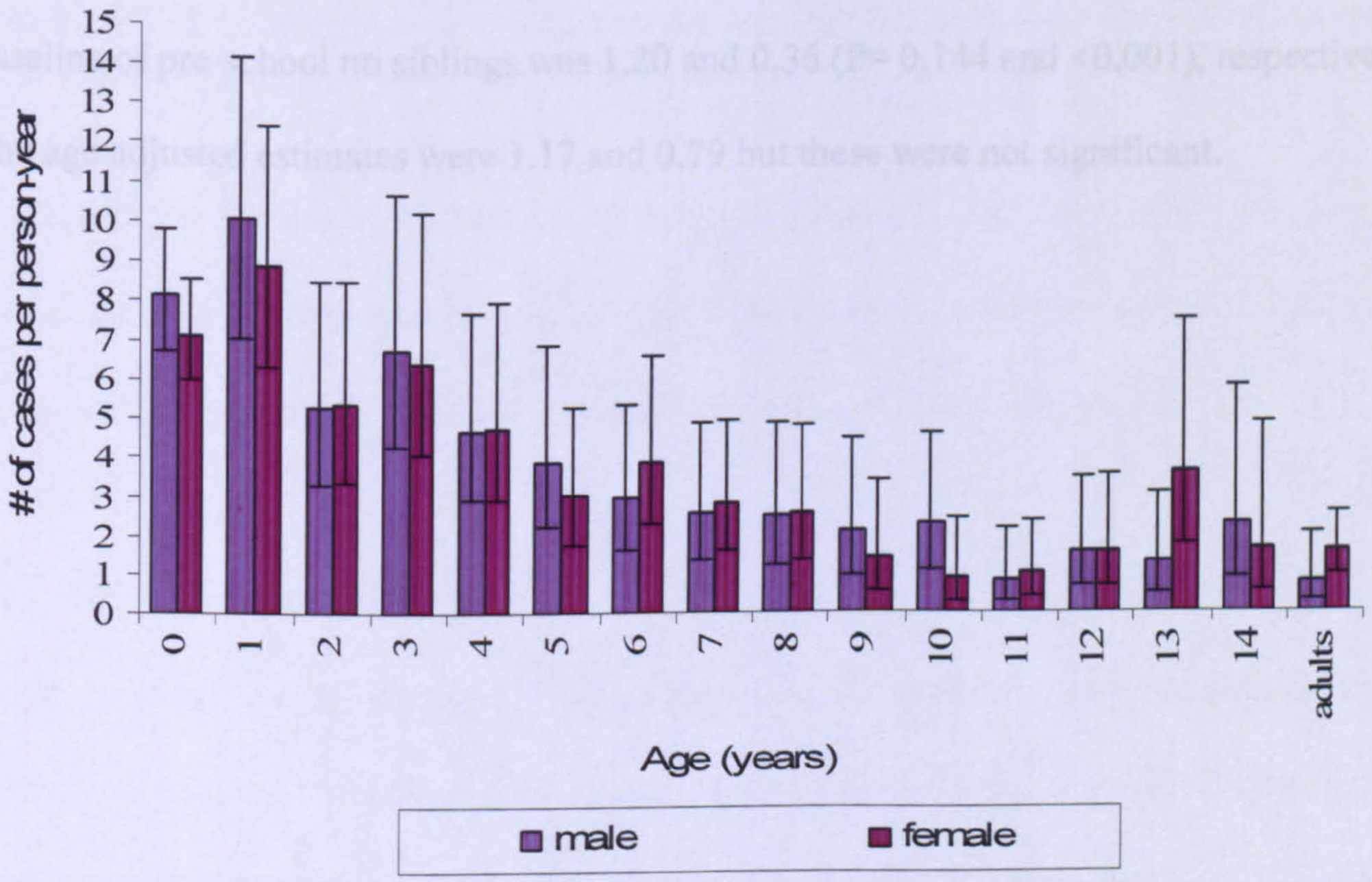


Figure 7.1. Incidence (95% CI) of all-cause respiratory illness among children in the household cohort by age and sex.

iii) Incidence of respiratory illness by child class

The relationship between school attendance and the incidence of respiratory infection was studied. Children were classified as detailed in the methods section of this Chapter. Figure 7.2 shows the age specific incidence rates in each of the three age classes. The influence of school children on respiratory illness among siblings is shown. For comparable ages, the rates tended to be higher in pre-school children with siblings who were going to school than those pre-school children without sibling(s) in school. Except for one age group, school children 3 years of age, the rates of infection in children attending school were lower than the rates for either class of pre-school children. The incidence rates are poorly established for 3 year-old school children because of limited data as few children of this age attend school. The crude IRR for pre-school with school siblings and for school siblings relative to baseline of pre-school no siblings was 1.20 and 0.36 (P= 0.144 and <0.001), respectively. The age adjusted estimates were 1.17 and 0.79 but these were not significant.

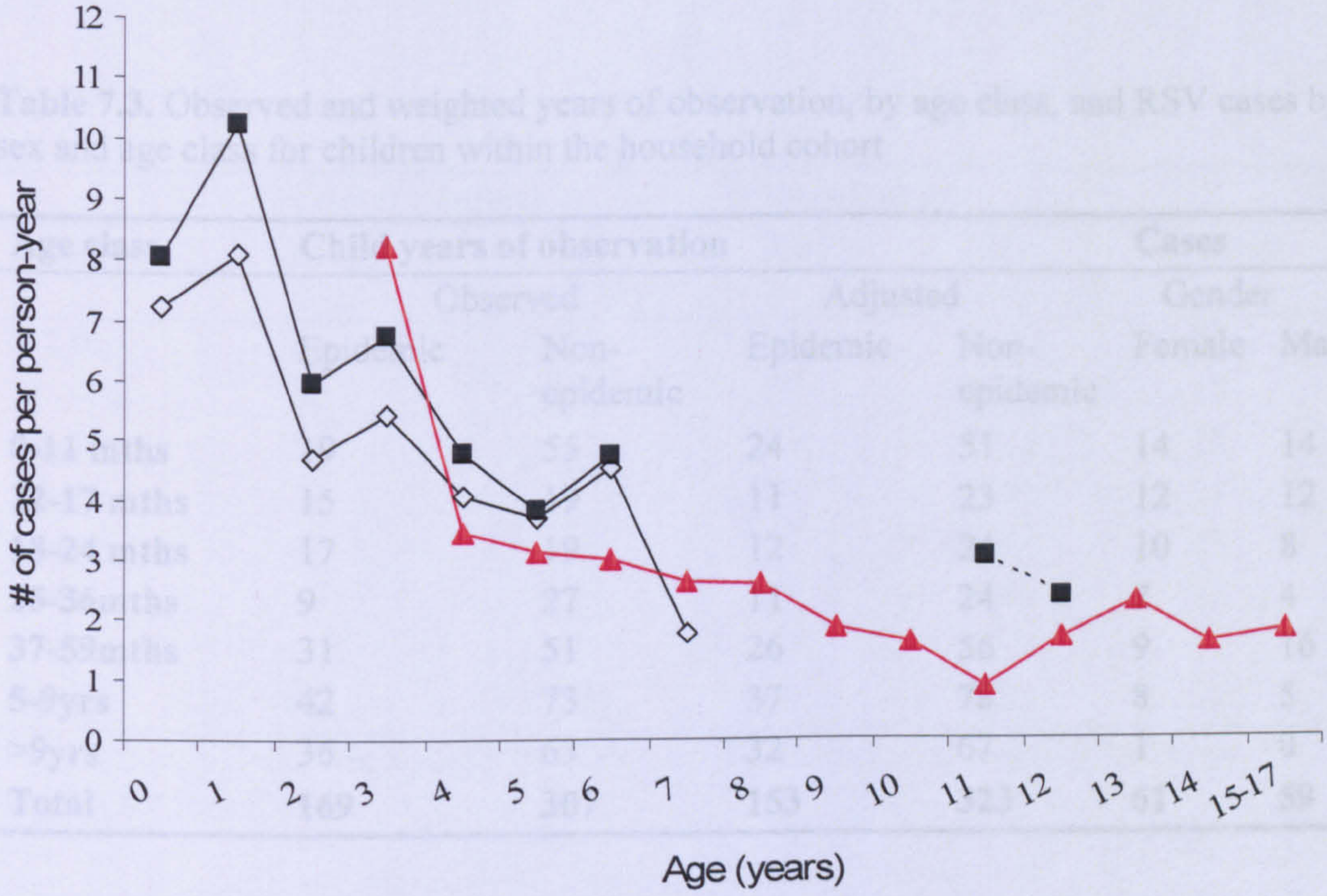


Figure 7.2. Incidence estimates of respiratory illness by age and child class: Pre-school children with no school child in the HH (open diamonds), pre-school child with sibling(s) in school (black boxes), school children (red triangles) and school-aged children not in school (black box dashed line). The 95% CI are given in Appendix P (p 353).

The importance of school children in increasing ARI rates in adult household members was demonstrated comparing rates in adults in households with and without school children (IRR=1.40) although this difference was not significant (Wald test, P=0.37).

7.4.2 RSV infection and disease

As only one adult had an identified clinical infection, these analyses were restricted to children (i.e. those < 15 years at recruitment). Details of child years of observation (cyo) and adjusted (weighted) cyo are summarized in Table 7.3 as are the cases of RSV infection and re-infection identified during the study by age and gender. In children 2 years of age or less there is no difference in numbers infected by gender. In children 3- < 5 years of age there were almost as many male children infected as females.

Table 7.3. Observed and weighted years of observation, by age class, and RSV cases by sex and age class for children within the household cohort

Age class	Child years of observation				Cases	
	Observed		Adjusted		Gender	
	Epidemic	Non-epidemic	Epidemic	Non-epidemic	Female	Male
0-11 mths	19	55	24	51	14	14
12-17 mths	15	19	11	23	12	12
18-24 mths	17	19	12	24	10	8
25-36mths	9	27	11	24	7	4
37-59mths	31	51	26	56	9	16
5-9yrs	42	73	37	78	8	5
>9yrs	36	63	32	67	1	0
Total	**169**	**307**	**153**	**323**	**61**	**59**

There were 22 repeat RSV infections (re-infections) identified during the study. As the history of infection was not known for the siblings in the household study, it was assumed that all RSV infections identified in children over two years of age were classified as re-infections and those infections in siblings 2 years of age or younger (5 children in total) were primary infections. Most of these are siblings born during the follow up period. Using this re-assignment, the number of primary and re-infections identified in this study by age class is shown in Table 7.4. Thus there were 54 primary infections identified in the household cohort and 66 re-infections. 59 % (39/66) of the re-infections were in children over 3 years of age.

Table 7.4. RSV disease risks after primary and re-infection stratified by age

	Primary Infection		**Re-infection**		
Age class	URTI	LRTI (risk %)	URTI	LRTI (risk %)	Total
0-11m	24	2 (8)	1	0 (0)	27
12-17m	12	4(25)	2	0 (0)	18
18-24m	5	1 (17)	9	2 (18)	17
25-36m	-	-	10	1 (9)	11
37-59m	-	-	25	0 (0)	25
5-9yrs	-	-	13	0 (0)	13
>9yrs	-	-	1	0 (0)	1
Total	**41**	**7 (15)**	**61**	**3 (5)**	**112**

Note: Fisher's Exact test by age: P= 0.202, for LRTI following primary infection; P =0.207, for re-infection.

The risk of developing disease after a primary infection and after re-infection was investigated. Data from twins (3 sets) was excluded from this analysis (total of 8 infections excluded). There was no significant difference in risk of disease by age after primary infection; 0-11m vs. 12+ months (RR=1.72, P=0.1414). There was borderline association (RR=0.41, p=0.055) between age and risk of disease following re-infection; 0-24 month vs. 25+ months. 15% of children <25 months getting an LRTI compared only 3% in older children (Table 7.4). However, there is little confidence in the age-stratified data as there are not many cases of LRTI. Thus an estimate for the total across age groups (more reliable) was made. The risk of disease following primary infection was higher (borderline significance) than after re-infection (RR= 0.32, P=0.069). There was little overlap in age classes of primary and secondary cases to enable a comparison. No difference in disease risk by sex was observed (RR =1.03, P=0.953). LRTIs were further subdivided to identify cases of severe RSV-LRTI. The risk of severe disease after primary infection was limited to children less than 18 months. There was no significant difference in risk of severe disease in the two age groups (Figure 7.3); 9.1% in children 0-11 months and 13% in children 12-

17 months (Fisher's Exact test, P= 1.000). There were no cases of severe LRTI resulting from re-infections.

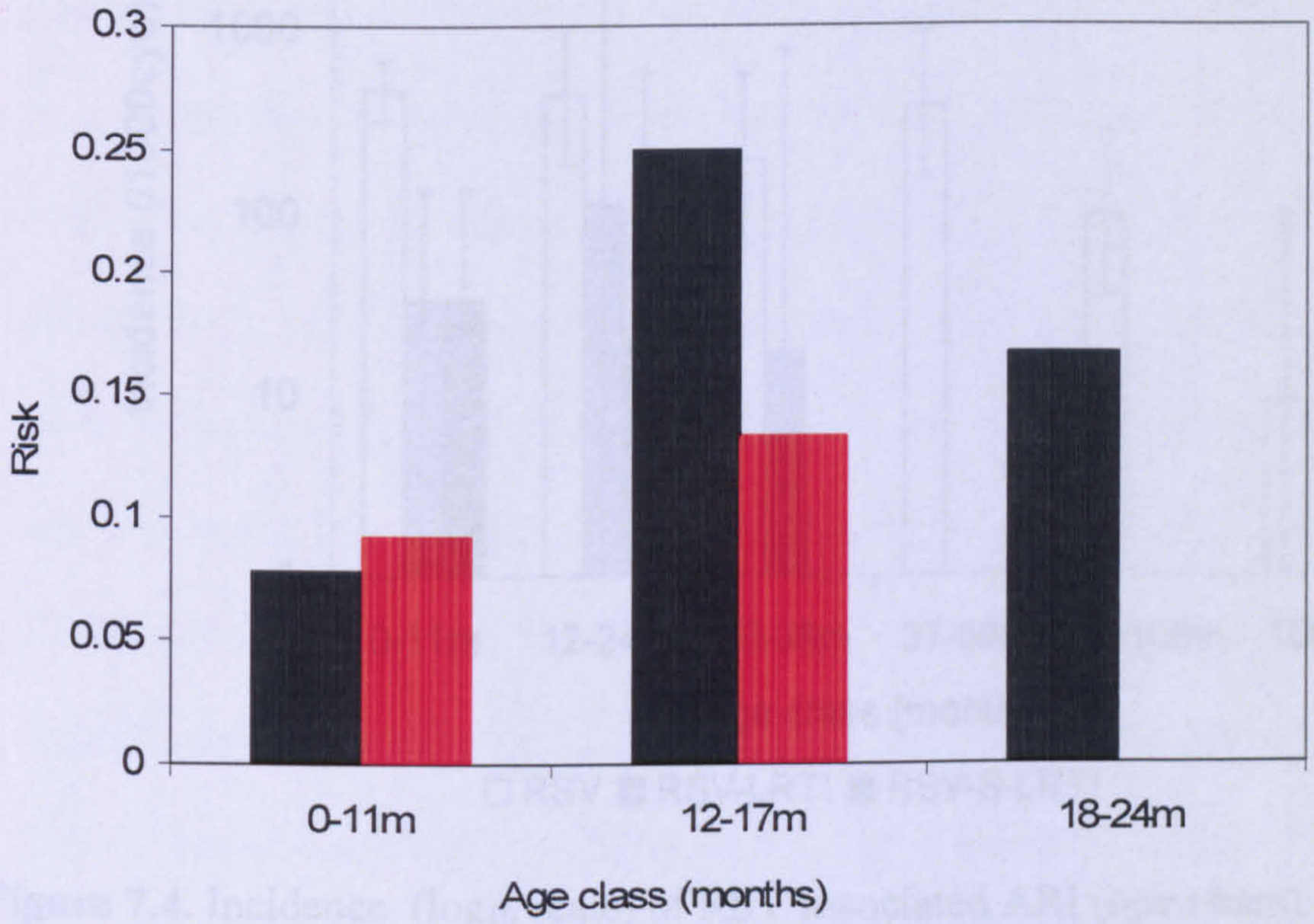


Figure 7.3. Proportion of RSV primary infections resulting in disease (Risk) stratified by age class. RSV-LRTI (black bars), severe RSV-LRTI (red bars).

7.4.3. Incidence of RSV infection

The incidence of RSV infection was 218 cases/1000 cyo (95% CI, 182- 264 cases/1000 cyo). The incidence of primary RSV infection was 476 cases/1000 cyo (95% CI, 361-630 cases/1000 cyo) and the incidence of re-infection was 147/1000 cyo (95% CI, 115-189 cases/1000 cyo).

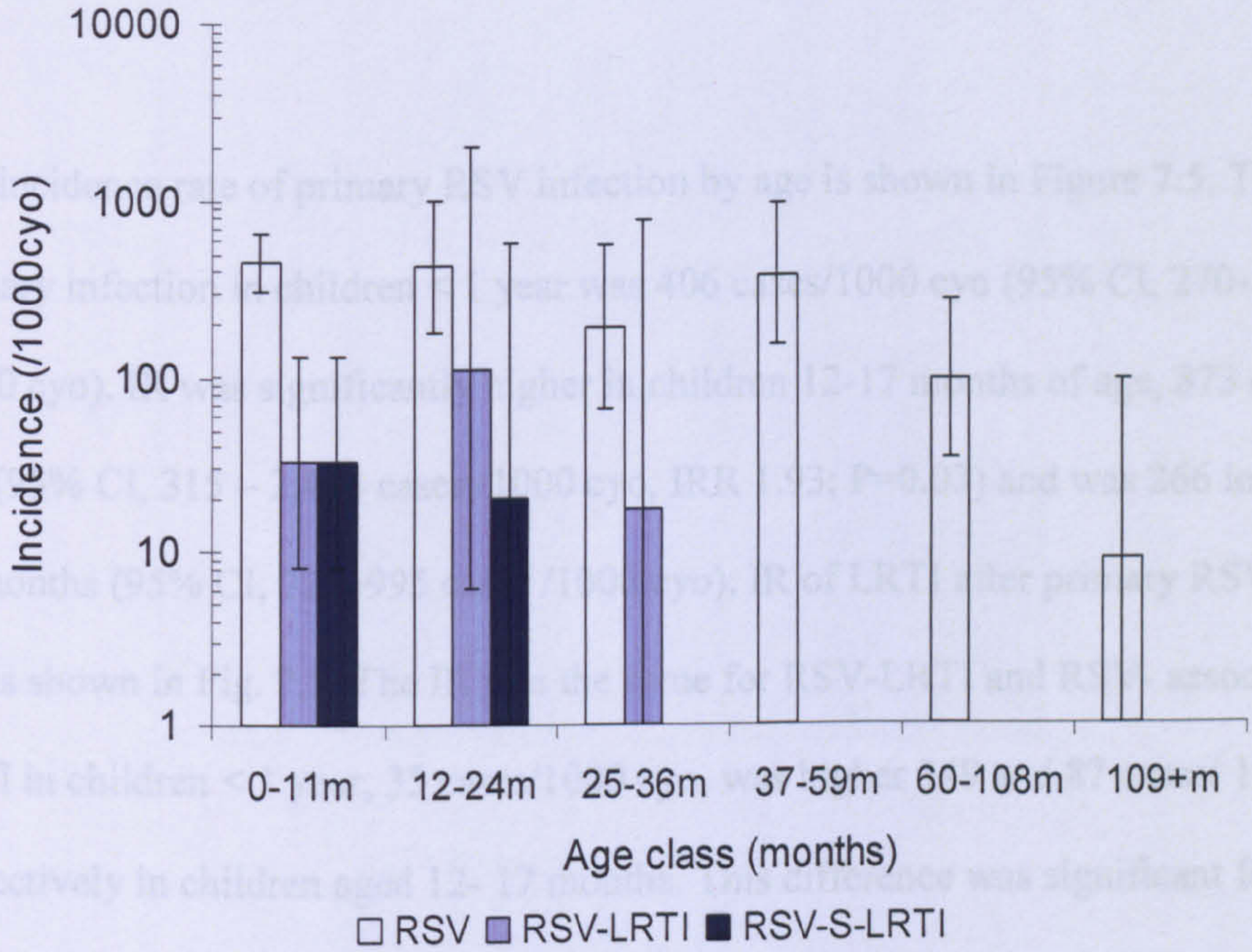


Figure 7.4. Incidence (log_{10} scale) of RSV associated ARI (open bars), LRTI (light blue bars) and severe LRTI (dark blue bars) by age class in the household cohort. Bars indicate 95%CI.

Figure 7.4 shows the incidence of RSV infection and disease in this study population. Estimates are inclusive or nested i.e. RSV-LRTI means RSV-LRTI and sLRTI. The actual incidence estimates are given in Appendix Q (p 355). Note that an RSV infection on its own must at least indicate an URTI. The estimates are inclusive or nested i.e. RSV-LRTI means cases of RSV LRTI or sLRTI. No appreciable decline was seen in infection rates of RSV until after 3 years of age where the rate of infection is significantly lower than for children 0-11 months (254 compared to 459/1000 cases)). LRTI occurred in children 3 years of age or lower while cases of severe disease were confined to children below 18 months. The IR of all-cause LRTI was 354 cases/1000cyo (95% CI, 282-444), that of RSV-LRTI was 22.5 cases/1000 (95% CI, 13 – 39 cases /1000 cyo). The IR of all-cause severe LRTI was 132 cases/1000 cyo (95% CI, 94-185), that of RSV- associated severe LRTI was 7.6 cases/1000(95% CI, 2.7–20.9 cases /1000 cyo).

The incidence rate of primary RSV infection by age is shown in Figure 7.5. The IR of primary infection in children < 1 year was 406 cases/1000 cyo (95% CI, 270- 612 cases /1000 cyo). IR was significantly higher in children 12-17 months of age, 873 cases/1000 cyo (95% CI, 315 – 2,414 cases /1000 cyo, IRR 1.93; P=0.03) and was 266 in children 18-24 months (95% CI, 72 – 995 cases /1000 cyo). IR of LRTI after primary RSV infection by age is shown in Fig. 7.5. The IR was the same for RSV-LRTI and RSV- associated severe LRTI in children < 1 year, 35 cases/1000 cyo, was higher 349 and 87 cases/ 1000 cyo, respectively in children aged 12- 17 months. This difference was significant for RSV-LRTI. The IR was 45 cases/1000 cyo (95% CI, 1- 1969 cases /1000 cyo) for RSV-LRTI in children 18-24 months with no cases of severe disease identified in this age group. There were no cases of hospital admission in the household cohort. One child died within a week of a positive RSV diagnosis (age 2 years 10 months with a mild LRTI) but the cause of death was unknown.

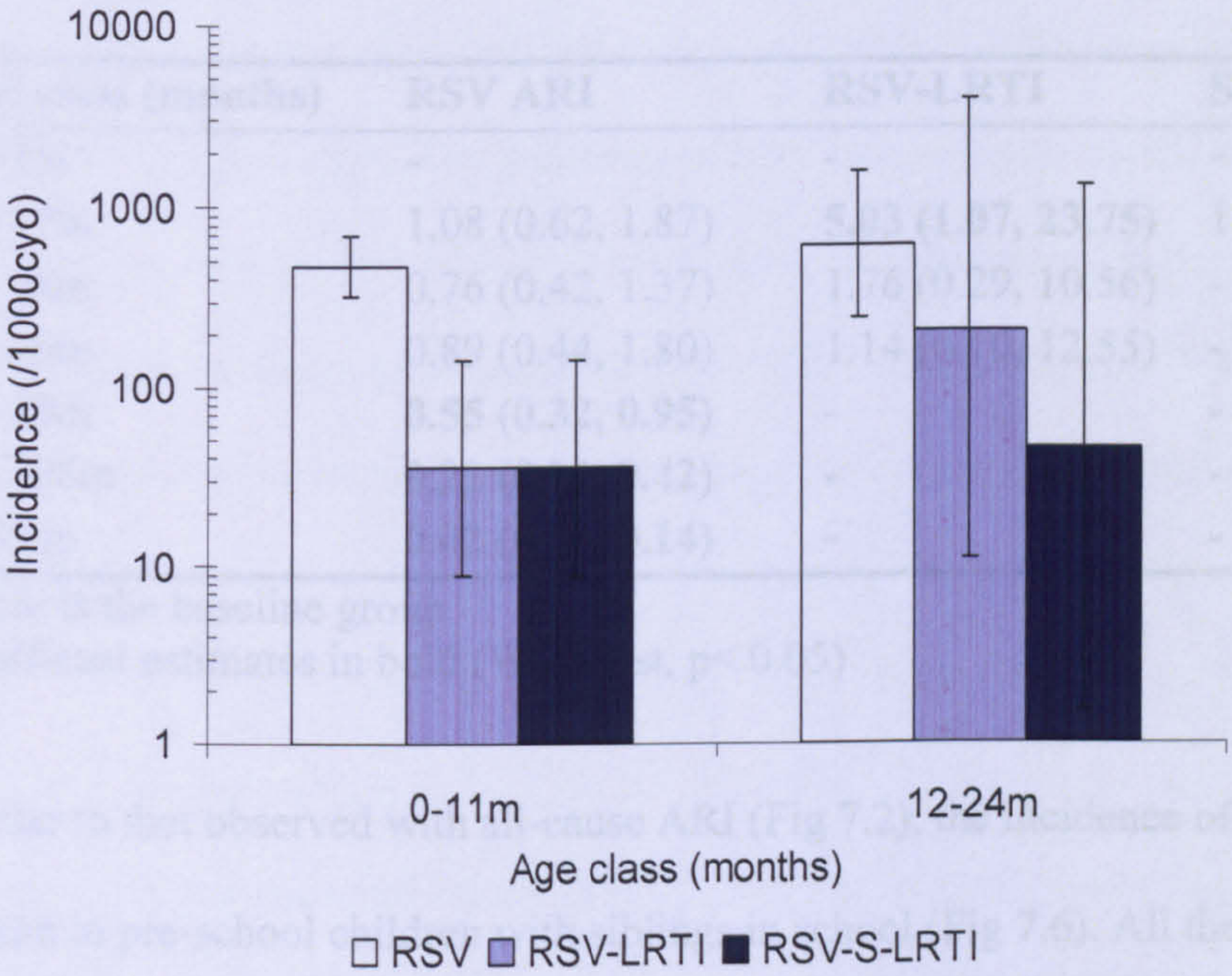


Figure 7.5. Incidence (log_{10} scale) of primary RSV associated ARI (open bars), disease (light blue bars) and severe disease (dark blue bars) by age class in the household cohort. Bars indicate 95%CI.

The age specific incidence rate ratios (IRR) of all RSV infections and disease are shown in Table 7.5. The comparison group is children 0-11 months of age. Adjustments for sex does not change the estimates significantly (lrtest, $p>0.05$). Incidence of infection declined with age significantly at age >3 years. There was only one child infected over the age of 10, a 14 year old girl with an RSV ARI. The IRR of RSV-LRTI in age class 12-17 months was 5 times that of children 0-11 months, this difference was significant (P=0.041). That of children in 18-24 months and 25-36 months was 1.8 times and 1.1 times that of the 0-11 month-old children, respectively (not significant). The IRR of RSV-associated severe LRTI in children 12-17 months was 1.25 times that of infants. This increased incidence was not significant. There was no case of RSV-associated severe LRTI in children older than 17 months.

Table 7.5. IRR of RSV-ARI, RSV-LRTI and RSV-associated severe LRTI by age class in the household cohort

Age class (months)	RSV ARI	RSV-LRTI	Severe RSV- LRTI
0-11m	-	-	-
12-17m	1.08 (0.62, 1.87)	**5.03 (1.07, 23.75)**	1.26 (0.18, 8.96)
18-24m	0.76 (0.42, 1.37)	1.76 (0.29, 10.56)	-
25-36m	0.89 (0.44, 1.80)	1.14 (0.10, 12.55)	-
37-59m	**0.55 (0.32, 0.95)**	-	-
60-108m	**0.22 (0.11, 0.42)**	-	-
109+m	**0.02 (0.00, 0.14)**	-	-

0-11m is the baseline group
Significant estimates in bold (Wald test, $p < 0.05$)

Similar to that observed with all-cause ARI (Fig 7.2), the incidence of RSV infection was highest in pre-school children with siblings in school (Fig 7.6). All the cases of RSV-LRTI identified were also in pre-school children with one or more siblings in school.

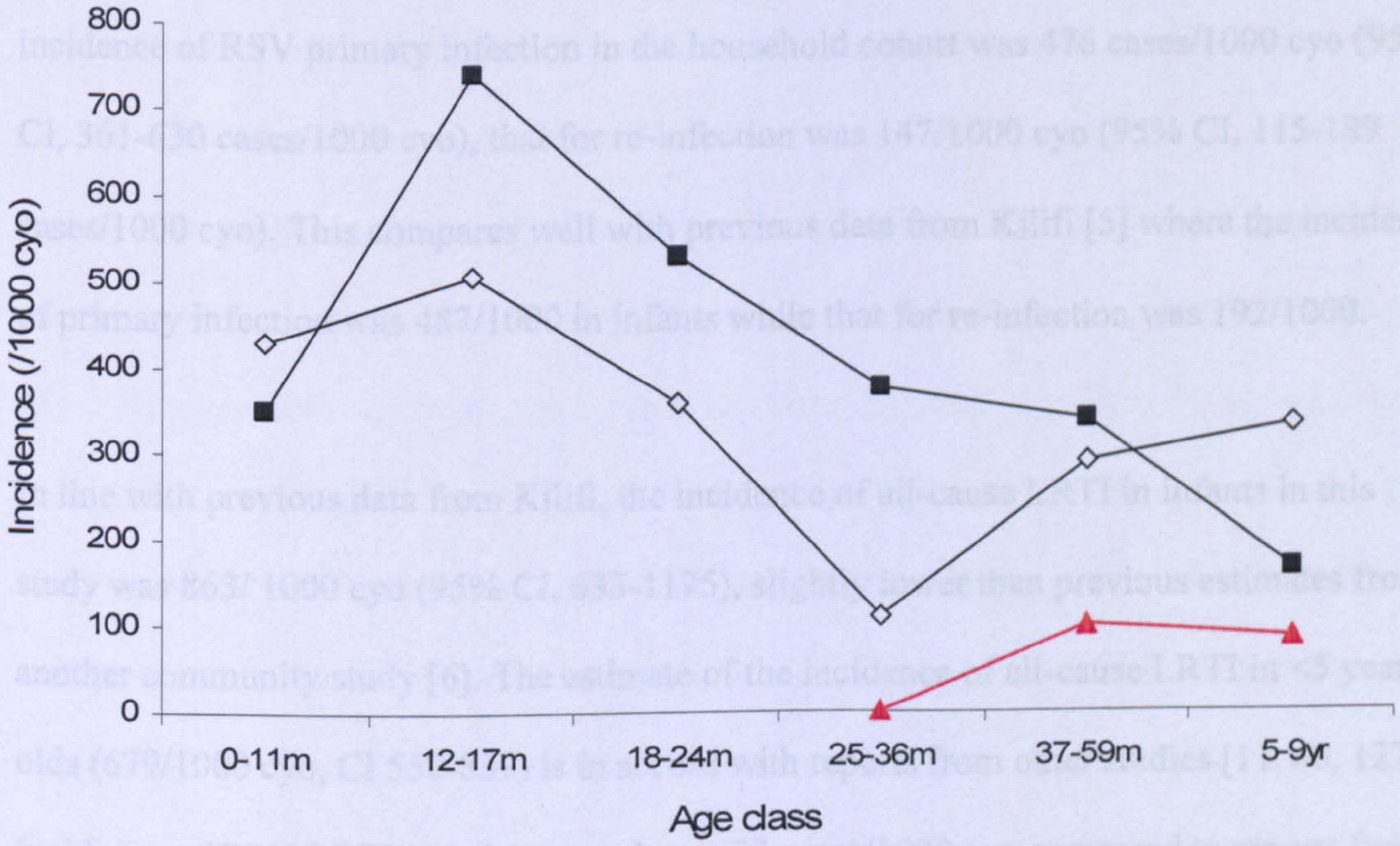


Figure 7.6. Incidence estimates of RSV infection by age and child class. The lines correspond to pre-school children with no school child (open diamonds), preschool child with sibling in school (black box) and school child (red triangles). 95% CI are provided in Appendix P (p 353).

The IRR of infection was stratified by age and gender. Differences by gender were not statistically significant. The incidence of RSV infection was compared within and between seasons for each quarter of each year of follow-up. The overall incidence of infection within each epidemic (468 cases/1000 cyo 95% CI 392-560) did not differ between epidemics with no transmission observed outside the main epidemic season (Wald test, $p>0.05$).

7.5 Discussion

This is the first household study of RSV from a Less Developed Country (LDC). During the study of 28 months duration and spanning two and a half epidemics, we detected an RSV infection in 31% of the children in the household cohort. 45% of the RSV infections were identified to be primary infections, 15% of the primary cases were LRTIs. The incidence of RSV primary infection in the household cohort was 476 cases/1000 cyo (95% CI, 361-630 cases/1000 cyo), that for re-infection was 147/1000 cyo (95% CI, 115-189 cases/1000 cyo). This compares well with previous data from Kilifi [5] where the incidence of primary infection was 487/1000 in infants while that for re-infection was 192/1000.

In line with previous data from Kilifi, the incidence of all-cause LRTI in infants in this study was 863/ 1000 cyo (95% CI, 633-1175), slightly lower than previous estimates from another community study [6]. The estimate of the incidence of all-cause LRTI in <5 year olds (679/1000 cyo, CI 550-839) is in accord with reports from other studies [11, 40, 127]. Incidence of RSV-LRTI in infants was lower 33 cases/1000 cyo compared to reports from other community studies from developing countries, 100-220 cases/1000 cyo [5, 6, 11] but similar to that reported from the study from Indonesia [11]. However, the incidence of

RSV-LRTI in under 5s was 45 cases/1000 cyo, lower than in Nigeria but similar to that reported in Indonesia [11].

The data show high incidence of RSV infection (ARI) well into young childhood, with no appreciable decline until >60m of age. High rates of RSV infection beyond infancy is neither unusual nor epidemiologically unexpected. The study by Glezen *et al* [12] showed high rates of RSV in year 2 of life and reasonably high rates in year 3 and 4. Similar results of significant incidence rates of RSV ARI in older ages and even into adulthood have also previously been reported from the family study by Hall *et al* [15]. Mixing patterns in this community would suggest high rates of contact in large households and school attendance suitable for RSV transmission. However, the data also show a higher rate of RSV disease in the second year of life. This is unexpected and is probably influenced by methodological issues. Only two clinical infections were identified in the first epidemic as a direct result of timing of recruitment in relation to the first epidemic in the HH cohort (i.e. it occurred earlier than had been anticipated). Thus most first infections in infancy were found in children with an average age of 11m that is nearing the end of their first year of life. Many of the child years of observation in infancy would therefore have been at the tail end of the first epidemic when the infants were very young with highest levels of maternal antibody and thus would be afforded protection from disease. Thus the higher rate of disease in older children is an artifact of the temporal incidence and cohort age, which has greater impact on the incidence of disease estimates rather than the incidence of infection. Data from the full birth cohort data was able to adjust for this because of the two cohorts. Results from this do nevertheless still indicate that significant risk of disease does exist beyond the first year of life (Nokes et al, in press). Notably a quarter of all re-infections were associated with LRTI. An attempt was made to account for this study design in the calculation of incidence

estimates by the weighting procedure detailed herein. Although we have attempted to overcome the design problems by weighting admittedly data from the household study and specifically the disease results are probably still influenced by this problem.

This study demonstrates quite considerable RSV infection levels in families (i.e. re-infection rates) as has previously been described in studies from developed countries [15, 16, 50]. In this study the incidence of clinical infection was drastically reduced after nine years of age; 101 cases in children 5-9 years to 9 cases/1000 cyo in older children. In the Tecumseh study in the US the highest rates (determined by serology) were observed in children 5-9 years [16]. Several factors may explain this. It has been observed in this and other studies that the incidence of ARI decreases with age. It is probable that there were missed asymptomatic RSV infections or infections with very short symptomatic periods. Nevertheless, several studies have reported that asymptomatic carriage of RSV is uncommon [261, 262] but not impossible [50, 101]. In order to consider asymptomatic infection, samples would have to be taken even in the absence of symptoms but justifying this might prove difficult. In addition, studies have shown that rapid antigen detection tests for the diagnosis of RSV respiratory illness in older individuals and adults are less sensitive than for diagnosis in children [125, 241]. Thus it is likely, as a result of the diagnostic method, that some infections may have been missed in older individuals.

The incidence rates of all-cause ARI reported here are similar (7.1 per child-year) to that from other studies, 6.1 per child-year in children <5 years in Manila [127]. As expected, respiratory infections were more common among female adults than male adults; a likely result of the closeness and length of contact that mothers have with children relative to fathers. Pre-school children with siblings in school had higher infection rates than pre-

school children without siblings in school. Presumably the former group is exposed to infection from their brothers and sisters who are attending school. Several studies done in developed countries provide evidence that younger children acquire infection from school-aged children within the household [15, 16, 50, 145].

Epidemics were characterized by periods of fadeout. To elucidate the mechanism of RSV persistence the possibility of the persistence of infections in the inter-epidemic period was considered. No clinical infections were identified between epidemics. Again it is possible that sub-clinical infections were taking place but such information would only be available from serological data which is presently not available. Several factors have been suggested as possible triggers of epidemics [39, 49, 60] from climatic factors to behavioural changes to loss of immunity and geographical position. In this environment it is still not clear what triggers epidemics as the intervals are variable (see Fig2.3, KDH IP data) with no clear relationship to climatic indicators or known behavioural changes related to weather patterns. Continued surveillance may shed further light on this issue in the future.

Certain methodological issues need to be qualified. New ARIs were delineated if symptoms were present one week after the end of the previous illness episode. Given that the aetiology of all cause ARI is diverse, there is justification for using only one week. Separate from this, the first 14 days following an RSV infection were excluded from the denominator since one was considered to be at no risk during this period. This was not done for ARI. It is rational to do this for RSV where repeat infections in the same child are relatively infrequent. However, for infections that are very common (all cause ARI), this is not practical as most observational time could potentially be excluded e.g. at the extreme case where children are getting a new ARI very frequently, there would be no observational time left and an incidence estimate can thus not be made. RSV Incidence was calculated

using weighted observation time (discussed in the methods). To illustrate the significance of this weighting a comparison of weighted and un-weighted incidence estimates was carried out (Table 7.6). Observed differences in estimates indicate that un-weighted observation time would result in either over or under estimation of the age-stratified incidence. This was not done for all-cause ARI based on the reasoning that ARI is less seasonal and therefore the age bias seen with RSV epidemic and non-epidemic observation time was likely to be less important. When age classes are combined this weighting effect ceases to be as important (Appendix R, p 356)

Table 7.6. RSV incidence (/1000 cyo) estimates comparing weighted and un-weighted observation time, for the household cohort

Age class (months)	**Infection**		**RSV-LRTI**		**severe LRTI**	
	weighted	unweighted	weighted	unweighted	weighted	unweighted
0-11m	459	371	33	26	33	26
12-17m	495	689	165	229	41	57
18-24m	346	494	58	82	0	0
25-36m	409	309	37	28	0	0
37-59m	254	300	0	0	0	0
60-108m	100	113	0	0	0	0
109+m	9	10	0	0	0	0

In conclusion we have demostarted an important burden of RSV infection and associated disease in relation to age in a rural community in Kenya and in Africa. We have documented the risk of disease and severe disease resulting from both primary infections and re-infection an aspect that has not been comprehensively studied. Thus these data attend to the need for more information from developing countries. In subsequent chapters, we explore factors in this population that are associated with increased risk of RSV infection and RSV-LRTI.

Chapter Eight

Risk Factor Survey

8.1 Introduction

There are three possible avenues of control for respiratory disease: prevention through vaccination, case management and supportive treatment for those who are already infected and lastly the identification and removal or avoidance of risk factors [194, 263]. It may be desirable to implement one or more or all of the three control methods. However, which of these will be of most use in the resource poor settings will depend on numerous factors, but of particular concern in the case of viral associated respiratory disease is the near absence of effective vaccines. The use of therapy in the form of anti-viral agents is unlikely to be an affordable wide-scale option, and proper case-management such as the use of oxygen or cohorting to prevent nosocomial spread are beyond the resources of the typical District Hospital in the sub-Saharan setting. Thus while immunization and case management are, or could be, partly effective in ARI [264], it is possible that the long term solution in developing countries will depend on the control of risk factors. It has been estimated that up to 25% of moderate and severe ARI might be preventable through addressing certain factors, namely exposure to wood smoke, nutritional factors, all of which are known risk factors for all-cause ARI [166, 194].

Respiratory infectious disease accounts for much global morbidity and mortality [21, 22, 25, 265, 266], and a significant proportion of ARI morbidity is due to RSV [27, 28]. Despite 40 years of research on RSV, we are really not anywhere near to licensing a vaccine, and oxygen therapy, taken for granted in industrialized countries, for supportive therapy for RSV, is not affordable in most resource poor settings.

Risk factor studies on disease caused by RSV from developing countries are few [126, 194, 200] and also none as yet has investigated risk factors for possible non-hospitalised severe disease occurring in the community [5]. The birth cohort (Chapter 3) provides a frame work for the comparison of risk factors for mild RSV infection and disease as well as comparing risk factors for non-specific LRTI (defined in chapter 9) with those identified for RSV-LRTI.

8.2 Chapter Aims

This chapter describes in detail the methodological approach adopted to collect the household and individual level risk factor data for RSV infection and RSV-LRTI and for all cause LRTI. This was a cross-sectional survey conducted towards the end of longitudinal follow up study of a birth cohort. The study design, data collection procedures and data management processes are described. The statistical background for the methods used in the analysis of risk factors (Chapter 8) and some general results are described. Issues arising from the study design, the appropriateness of the questionnaire and the computation of the socio-economic asset index are discussed.

8.3 Survey

8.3.1 Study Design and Ethical Considerations

A total of 635 children recruited at birth were monitored over 2 and half years between 2002 and 2005 for episodes of RSV infection through a process of passive referral and active surveillance as described in Chapter 2. Within this framework, a cross sectional risk factor survey was carried out on families of all children from this cohort to investigate the association between exposure to putative risk factors for RSV specific infection/disease and all cause LRTI both in the individual and in the household. The survey was carried out between June 2004 and November 2004, that is at the end of the third epidemic season just

before follow up of cohort 1 ended and before the last year of follow up of cohort 2. All children in the main birth cohort were eligible to take part in the survey as long as they were still being followed up at the time of the survey. At the beginning of the interview visit, the purpose of the study was explained to the parents or guardians and verbal consent sought before the interview commenced. The field worker conducted the interview, reading out and explaining the questions to the respondent and then filling in the questionnaire. Ethical permission was given within the main proposal by the Kenya Medical Research Institute/National Ethical Review Committee and Coventry Research Ethics Committee, UK.

8.3.2 Questionnaires

A questionnaire (Appendix S, p 358) was filled out for all the children enrolled. The questionnaire addressed household characteristics, demographic, socio-economic and environmental factors. These were fashioned on the basis of previous work done on risk factors for RSV and other risk factor surveys conducted in sub-Saharan Africa [167, 175, 194, 267-269]. The questions on socio-economic status and household characteristics were used to classify households into different socio-economic classes based on an asset index described herein.

A pilot risk factor survey was initially conducted in collaboration with a student intern (Nicodemus Kisengese). The first pilot study was mainly used as a tool for accessing the appropriateness of questions included. Results from this pilot study were used to modify and inform the main survey questionnaire with certain questions being weeded out or modified and new questions included as necessary. It was also used to design an appropriate database to facilitate data entry. Most of the questions were multiple choice

with one or two applicable answers being marked with a circle. Several questions required a number to be entered into the box (e.g. how many children live in your household?) The question on occupation of the major income provider (MIP) was the only one to require a written answer. Each field worker was provided with a rubric (Appendix T, p 364) in English giving details of the questions in the questionnaire as well as details on what sort of responses were expected for each question. A field worker training session was subsequently conducted. The questionnaire was in English but questions were translated into the local Giriama or Kiswahili language at the time of the interview.

8.3.3 Data Management

Data from the questionnaires were double entered by two independent data entry clerks onto a FileMaker (FileMaker Pro Developer Version 5.5) database with several internal consistency checks. The two data sets were then compared using a specially designed verification program in FoxPro (Version 6) that generates a text file document detailing the discrepancies, which were resolved by consulting the original paper questionnaires.

8.3.4 Description of data set

Longitudinal data on infection history from active and passive surveillance was combined with cross-sectional data from the risk factor survey. The dataset consisted of individual subjects with observations recording the start and end of each observation period and the infection status at the end of that observation period resulting in survival time data. Each child had multiple record visit data over the follow-up period. A total of 529 questionnaires were completed including 60 from infants recruited form IP or OP. Information on household and individual demographic and social characteristics was available from the questionnaire data. The outcome variables were (i) LRTI (ii) RSV-ARI and (iii) RSV-LRTI, all identified by clinical surveillance methods using modified WHO criteria as

previously described (Sec 7.3, Chapter 7), and in the specific case of RSV identified by antigen assay (IFAT) of nasal washings. These three outcomes will be referred to as all-cause LRTI, RSV infection (RSVI) and RSV disease (RSV-LRTI) respectively in the following text of this and the next chapter.

8.3.5 Response rate

Out of a total of 635 birth cohort children who had at least one visit, data on risk factors was obtained from 469 children (74%). Most (75%) of the 166 children for whom no data was available had dropped out of the study before the start date of the risk factor survey. 469 children were used in this analysis with 60 children being excluded because they were recruited into the cohort from either the IP or OP and thus did not have records from birth.

8.4 Analysis

The analysis used for generation of the results presented in Chapter 8 are presented here.

8.4.1 Descriptive analysis

Standard descriptive statistical analyses (means, frequency tabulations) were conducted to explore the dataset. The distribution of each of the variable was examined to find outliers and to get a good understanding of the characteristics of the study population with respect to the risk factors. Single variables were investigated in relation to RSV infection/disease and all-cause LRTI incidence. Data reduction procedures were then carried out. Only variables with substantial variability within categories were included. Variables with categories that had small numbers of observations were combined. Categories were created after complete exploration of variables to ensure no effect was masked and check for non-linearity. Correlation and interaction between variables was explored creating new

variables that best captured the essence of the data. The means, ranges and distributions of these variables are in Appendix U with a description of social and demographic characteristics of the population can be found in Table 7.4.

8.4.2 Creation of an asset based wealth index to estimate household socio-economic status

In most African settings information on income and expenditure is difficult to collect because most people engage in subsistence farming, barter or petty trade and therefore information on income and expenditure is often unreliable. In the absence of information on household income and expenditure several alternative methods have been developed that rely on assets owned and housing quality as well as occupations of household members. In this study I used data on asset indicators or asset variables to construct an asset index based on the method described by [270]. This asset index does not reflect current consumption; rather it is viewed as a proxy for a household's long running economic status. This is done by creating a proxy for wealth by constructing a linear index from asset ownership indicators, using principal components analysis (PCA) to derive weights. Principal components analysis is a statistical technique for data reduction and provides a method of identifying those factors that most contribute to the variance of each composite wealth index. The objective is to find unit-length linear combinations of variables that contain the most information (greatest variance) and the weighting of the variable is based on the value of its principal component [270-272]. The first component (PC1) gives the minimum perpendicular distance between it and all the data points; minimising the perpendicular error is the same as maximizing the variation in the predicted i.e. the PC. Thus the PC1 maximizes the differences between households and is typically assumed to be the asset index. The result obtained from the first principal component is usually used to develop the asset index A_j based on the following formula:

$$A_j = f_1 * (a_{j1} - a_1) / (s_1) + \ldots\ldots + f_N * (a_{jN} - a_N) / (s_N)$$

In the above formula:

- f_1 is the 'scoring factor' for the first asset as determined by the procedure;
- aj_1 is the jth household's value for the first asset; and
- a_1 and s_1 are the mean and standard deviation of the first asset variable over all households [270].

i) Construction and validity of the asset index

The classification was based on the computation of asset indices that take into account a set of asset-based and health-related variables for each household:

- whether the household has electricity, radio, a television or video, any bicycles, any motorcycles, or a telephone (each coded as 1 = Yes, 0 = No);
- the main household source of drinking water (5 categories);
- the main type of cooking fuel (three categories).
- the main type of toilet facility used by the household (three categories);
- the main type of housing material in the household (two categories); and
- the home ownership (three categories).

The percentage of the variance explained by the first principal component is 24%, Table 8.1. The first eigenvalue is 6.096; the second eigenvalue is 2.36.

Table 8.1 Eigenvalues from the PCA. Showing the variance explained by the first 13 components.

Component	Eigenvalue	Difference	Proportion	Cumulative
Comp1	6.09621	3.73503	0.2438	0.2438
Comp2	2.36118	0.390062	0.0944	0.3383
Comp3	1.97111	0.199366	0.0788	0.4171
Comp4	1.77175	0.098659	0.0709	0.488
Comp5	1.67309	0.309788	0.0669	0.5549
Comp6	1.3633	0.160906	0.0545	0.6095
Comp7	1.2024	0.094942	0.0481	0.6576
Comp8	1.10745	0.092685	0.0443	0.7019
Comp9	1.01477	0.042767	0.0406	0.7425
Comp10	0.972002	0.147344	0.0389	0.7813
Comp11	0.824659	0.037738	0.033	0.8143
Comp12	0.78692	0.064033	0.0315	0.8458
Comp13	0.722888	0.060349	0.0289	0.8747

The first component has negative loadings for having no toilet, house walls made of mud, using fire wood as the main cooking fuel, owning the house in which you live, obtaining water from a well that is open to the public and owning an animal (cows, goats or sheep) thus can be interpreted as distinguishing the "poor" households from those that are better off households, Table 8.2. Owning a house usually implies living in your rural homestead and thus entails very little financial input (as most of them are makuti (palm leaves) thatched, mud-walled houses).

Table 8.2. First six components from the principal components analysis.

Variable	Comp1	Comp2	Comp3	Comp4	Comp5	Comp6
phone	0.2169	0.2564	0.0248	0.183	-0.0364	0.0734
radio	0.1282	0.1701	0.1863	0.2055	0.0286	0.1963
Motor bike	0.0079	0.0950	0.0985	0.1516	-0.0768	0.0289
TV or video	0.2191	0.2477	-0.0852	0.1395	0.0486	0.1604
bicycle	0.0378	0.2250	0.2480	0.2616	-0.0571	0.1677
Electricity	0.2529	0.1139	-0.1901	-0.0145	0.1123	-0.0277
Cows	-0.0066	0.2124	0.1553	0.1404	-0.2118	-0.2150
Goats	-0.0213	0.2894	0.4463	-0.4056	0.1541	-0.0478
Sheep	-0.0044	0.2327	0.4106	-0.4572	0.2136	-0.0354
Block walled house	0.3202	0.0379	0.1082	-0.0952	-0.0808	0.1645
Mud walled house	-0.3202	-0.0379	-0.1082	0.0952	0.0808	-0.1645
Owned house	-0.2797	0.2662	0.0504	0.2371	-0.0201	0.1803
Rented house	0.2629	-0.2775	-0.0446	-0.2369	0.0340	-0.0709
Not rented not owned	0.0752	-0.0097	-0.0197	-0.0318	-0.0288	-0.2722
Flush toilet	0.1836	0.2752	-0.2713	0.0077	0.2148	0.1273
Latrine	0.1333	-0.3571	0.3539	0.1239	-0.2552	0.1661
No toilet	-0.2313	0.2291	-0.2278	-0.1320	0.1545	-0.2366
Main fuel firewood	-0.3228	0.1336	0.0484	0.0953	-0.0643	-0.0038
Main fuel charcoal	0.3021	-0.1179	-0.0754	-0.0981	0.0741	-0.0848
Main fuel non biomass	0.0668	-0.1027	0.1165	0.0364	-0.0608	0.1865
Dw well	-0.0848	0.1206	-0.2025	-0.3399	-0.4656	0.286
Dw piped	0.0944	-0.1428	0.2207	0.3341	0.4718	-0.2463
Ws owned	0.2310	0.1873	-0.2355	0.0057	0.1698	0.1324
Ws public	-0.2636	-0.2208	0.0682	-0.0628	0.2365	0.3496
Ws shared	0.1437	0.1437	0.0877	0.0662	-0.4140	-0.5110

Table 8.3 reports the scoring factors from the PCA of 25 variables. Scoring factors are the "weights" assigned to each variable (normalized by its mean and standard deviation) in the linear combination of the variables that constitute the first principal component. When using a correlation matrix the PCs are in standardized units (mean 0, variance 1). In such a case the scoring coefficients/ factors and the eigenvectors (PCs) will be the same.

Table 8.3. Scoring factors for variables entering the computation of the first four principal components.

Variables*					Means		
	scoring factors	mean	std.dev. (sd)	scoring factor/sd	Poorest 34%	Middle 33%	Richest 33%
own phone	0.217	0.225	0.418	0.519	0.017	0.155	0.509
own radio	0.128	0.664	0.473	0.271	0.406	0.776	0.817
own motorbike	0.008	0.009	0.097	0.082	0.000	0.011	0.017
own a TV or video	0.219	0.134	0.341	0.642	0.000	0.063	0.343
own bicycle	0.038	0.484	0.500	0.076	0.311	0.615	0.531
electricity in home	0.253	0.074	0.262	0.967	0.000	0.000	0.223
Number of cows owned	-0.007	0.76	2.169	-0.003	0.739	0.759	0.783
Number of goats owned	-0.021	3.155	9.839	-0.002	2.750	4.150	2.580
Number of sheep owned	-0.004	0.386	6.593	-0.001	0.028	1.011	0.131
house made of block wall	0.320	0.363	0.481	0.665	0.000	0.195	0.903
house made of mud wall	-0.320	0.637	0.481	-0.665	1.000	0.805	0.097
owns house you live in	-0.280	0.784	0.412	-0.680	1.000	0.931	0.417
rent house you live in	0.263	0.185	0.389	0.676	0.000	0.046	0.514
not rented not owned	0.075	0.03	0.171	0.439	0.000	0.023	0.069
HH owns flush toilet	0.184	0.062	0.242	0.758	0.000	0.006	0.183
HH owns latrine	0.133	0.594	0.492	0.271	0.646	0.206	0.800
HH has no toilet	-0.231	0.344	0.476	-0.486	0.794	0.207	0.017
Cooking fuel is firewood	-0.323	0.745	0.436	-0.740	1.000	0.920	0.309
Cooking fuel is charcoal	0.302	0.242	0.429	0.705	0.000	0.086	0.646
Cooking fuel gas or kerosene	0.067	0.03	0.171	0.390	0.000	0.017	0.074
drink water from well	-0.085	0.059	0.235	-0.361	0.133	0.023	0.017
drink water from piped source	0.094	0.896	0.306	0.309	0.778	0.948	0.966
water source owned	0.231	0.083	0.276	0.836	0.000	0.006	0.246
water source public	-0.264	0.775	0.418	-0.631	0.978	0.885	0.457
water source shared	0.144	0.147	0.355	0.405	0.022	0.109	0.314
Economic Status Index					-2.16	-0.82	3.03

*The means, ranges and distributions of these variables are in Appendix U (pg 367).

The index is essentially the sum of included variables, weighted by the elements of the first eigenvector. Because majority of the asset variables are binary (0/1), the assigned weights have an easy interpretation: a move from 0 to 1 changes the index by f_{1i}/s_i (column 4). An

asset that doesn't vary across households gets a weight of zero for instances as it explains none of the variation across households.

The difference in the average index between the poorest and the middle group is 1.34 units. One example of a combination of assets that would result in this difference is owning a TV/ video (0.642), using charcoal as your main cooking fuel (0.705) and owning a cow(s) (-0.003). The average asset index in richest group is 3.85 units higher than for the middle group. This difference is equivalent to owning a radio (0.271), TV or video (0.642) and phone (0.519), having electricity (0.967) and flush toilet (0.758), using charcoal as cooking fuel (0.705) and owning a cow(s) (-0.003) and goat(s) (-0.002).

Table 8.3 shows a consistent pattern in the asset variables used in the model. The assets that are likely to be owned by the better-off households have a positive value, which increases the household's asset index. In contrast, those that characterise poor households (e.g. no toilet or mud wall and an open well) have the expected negative value, which results in a decreased asset index. A household that owns a flush toilet has an asset index higher by 0.76 than one that does not; owing a radio raises a household's asset index by 0.27 units; drinking water from a well lowers the asset index by 0.36 [270].

Households were sorted by the asset index and grouped into percentiles. For expository convenience the top 33% are referred to as "rich" the next 33% as "middle" and the bottom 34% as "poor". This classification does not, and is not intended to, follow any of the standard definitions of poverty. The last three columns of Table 7.3 compare the average ownership of each asset across different types of households. Large differences are observed across groups showing the internal coherence of the asset index, radio ownership

is 40% for the poor versus 82% for the rich. Along the same line, houses with block wall are 0% for the poor and 90% for the rich. Households with no toilet are 79% in the poor compared to 1.7% in the rich households. Certain variables don't vary across the groupings (such as number of cows) while interestingly other variables have a non-linear pattern (such as number of goats and sheep or bicycle ownership).

8.4.3 Calculation of anthropometric indicators based on WHO child growth standards

Anthropometric measurements were obtained at birth and at three month intervals thereafter. We used a WHO macro in STATA (igrowup_STATA macro [273]) to calculate z-scores for three anthropometric indicators, weight-for-age, length/height-for-age, weight-for-length/height (Appendix V, p 368). Z-score is the deviation from the median of a reference population (the WHO Child Growth Standards) expressed in multiples of the standard deviation [274]. Any extreme (i.e. biologically implausible) z-scores were flagged up for each indicator according to the following system: weight-for-age z-scores of < -6 or > 5, length/height-for-age z-scores of < -6 or > 6 and weight-for-length/height z-scores of < -5 or > 5. These were marked as missing and were excluded.

For children aged below 24 months (< 731 days) and measured standing, the macro converts the height to recumbent length by adding 0.7 cm; and for children aged 24 months or above who are measured in recumbent position, the macro converts the length to standing height by subtracting 0.7 cm. In other words, all the z-scores for children below 24 months are length-based, and height-based otherwise. The macro imputes any missing data on whether measurement was of recumbent length or for standing height according to the following algorithm: a) If age is given, then it is recumbent length if the child's age is below 24 months (< 731 days), and standing height if the child's age is 24 months or above.

b) If age was missing, then it was assumed to be recumbent length if length/height < 87 cm and standing height if length/height ≥ 87 cm (Appendix V).

8.4.4 Regression models for cohort studies

In longitudinal studies risk factors and exposures are not constant and change with time and this variation must be taken into account. This is usually done by breaking the individual follow-up times into segments specific to the time changing variable and is known as Lexis expansion. Within each segment the rate is assumed to be constant. The individual follow-up time is split into bands of the time changing variable and then the Poisson regression is used [275]. The simplest model for cohort studies is the Poisson model.

Cox's method is very similar to the Poisson regression but is based on a much finer subdivision of time [275]. Cox regression is a Poisson model with very narrow time bands such that each time band only has one event. This is also known as a proportional hazard model and is based on the proportional hazard assumption which states that given two observations with different values for the independent variable, the ratio of the hazard functions for those two observations does not depend on (constant over) time. Cox regression is used if the hazard rate is thought to vary rapidly with time and also to evaluate the effects of explanatory variables or covariates on the hazard rate, i.e. expressed as a function of both the time and the covariate $\lambda(t,x)$ where λ is the hazard rate, t is time and x is the covariate(s). The hazard rate is modeled to represent dependence on variables recorded for each subject as well as on time. The form of the baseline hazard is not specified, and the model is therefore nonparametric with respect to time but parametric in terms of covariates [276].

The Cox model only operates on times when failures actually occur. Consequently in Cox regression the results are based on forming, at each failure time, the risk set and then maximizing the conditional probability of failure [275, 277]. The risk set is the collection of subjects who are at risk of failure at the time of the event. Different time scales can be used to compute risk sets. With a Cox model the choice of the time scale (i.e. time since entry into study or time as age or calendar date) determines the composition of the risk sets associated to each failure [277]and results produced are critically dependent on the choice of timescale. Time should be interpreted in the way that is most appropriate for the analysis to be carried out [275], i.e. the time scale with the strongest relationship to the failure rate, and in this case calendar time, will be used. This eliminated the potential confounding effect of seasonality as the resulting risk sets were defined at the calendar date of each event (children under the same level of exposure) (Figure 8.1).

When subjects are tied (i.e. failures that occur at the same time) and the exact ordering of failure is unclear, the partial likelihood needs to be modified. The Breslow approximation used in this analysis to deal with tied failures states that since the order of failure is not known, the largest risk set for each of the tied failure events is used. It is an approximation to the exact marginal likelihood, in which the likelihood reflects the marginal probability that the tied failure events occurred before the non-failure events in the risk pool, and the order in which they occurred is not important [278].

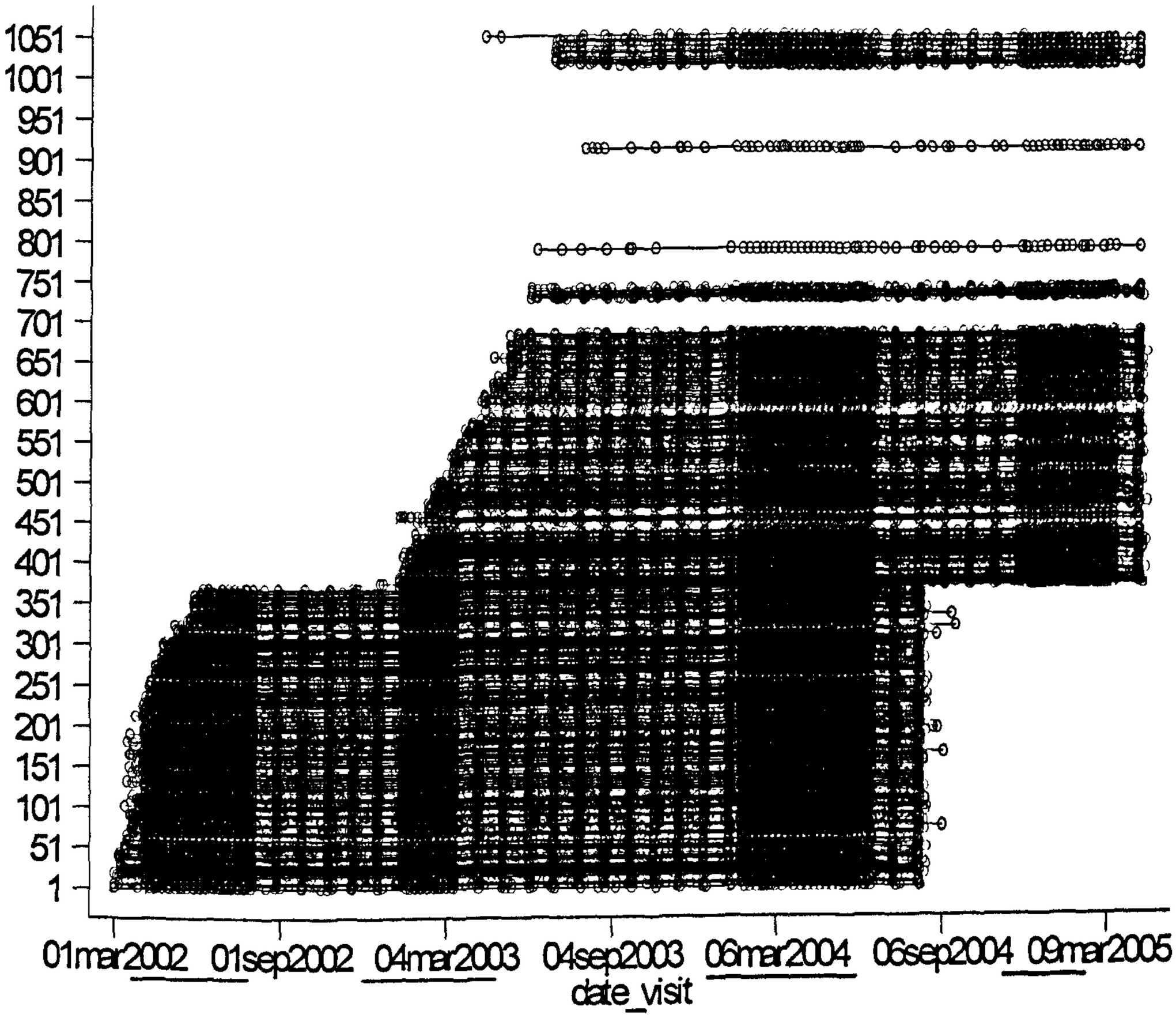


Figure 8.1. Follow up of cohort children by calendar time. Lines indicate individual follow-up of children. Time starts on date first study subject visited and ends on last visit date of follow up. The red lines at the bottom of the graph show the four epidemics that occurred during the study. These were between 12/03/02 – 01/07/02, 03/12/02 – 15/04/03, 08/01/04 – 02/06/04 and the last between 11/1104 – 18/02/05.

Time Dependent Variables

These are handled automatically by the Cox model in STATA. Since the risk set is indexed by the failure time associated with it, the model simply takes the value of the covariate at that time [279]. The premise is that the instant the value of the variable changes the hazard of risk changes [277]. In STATA when the time-varying nature of the covariate(s) is not spelled out in multiple observations per subject, the time varying covariates are specified using the tvc option [278] used when one wants a covariate to change as a continuous function of time, i.e.

time-varying covariate = (constant covariate in my data) * f(t)

where f(t) is a function of time, with f(t) = t by default and settable via the option *texp ()* in STATA [279] .

Clustering of Data

For clustered data, i.e. repeated measurements within individuals, a shared frailty model should be used. This is the survival data analog to regression models with random effects [277]. A frailty is a latent random effect that enters multiplicatively on the hazard function. A Cox model with shared frailty is therefore simply a random effects Cox model used to model within group correlation. Observations within a group/individual are correlated because they share the same frailty and this extent of correlation is measured by the parameter θ, and when zero the Cox shared frailty model simply reduces to the standard Cox model [278]. An alternative to fitting a shared frailty model is to fit the standard Cox model and adjust the standard errors of the estimated parameters to account for possible correlation using the cluster option in STATA. Both methods were used in the analysis as detailed in Chapter 8.

Testing the Proportional Hazard Assumption

One of the main assumptions of the Cox proportional hazard model is proportionality. Shoenfeld and scaled Schoenfeld residuals were used to test the proportionality assumption. This test for non-zero slope is the generalized linear regression of the scaled Schoenfeld residuals on functions of time. A non-zero slope is an indication of a violation of the proportional hazard assumption [280]. The regression graph was assessed using the *stphtest* command in STATA and a test of non-zero slope carried out. Log-log plots were also used

to assess the relationship, in which parallel lines indicate that predictors do not violate the proportionality assumption.

Assessment of Model Fit

The goodness of fit of the model can be evaluated using the Cox-Snell (cs) residuals. If the model fits the data well then the true cumulative hazard function conditional on the covariate vector has an exponential distribution with a hazard rate of one [277]. With multiple observations per subject, a residual that is subject specific, in this case summed over the observations within the subject is required: cumulative Cox-Snell residuals [277]. Under a good-fitting model, the cumulative hazard of the Cox-Snell residuals should be a straight 45° line [277, 281]. The model fit can be verified by estimating the empirical Nelson-Aalen cumulative hazard function, with the Cox-Snell residuals as the time variable along with the data's original censoring variable.

8.4.5 Univariate Analysis

First the data was declared to be survival-time data using the *stset* (st-survival time) command in STATA. Initial univariate analysis was carried out using Kaplan Meier (KM) curves and log rank tests [282]. KM curves were used to assess proportionality. Tests of equality of survival distribution across strata were performed using the log-rank test for categorical variables. Univariate Cox proportional hazard regression was used for continuous variables. These were done to explore whether or not to include the risk factor in the final model. The risk factor was included if either test gave a P-value of 0.25 or less. This way, only those variables that had an unconditional association with the outcome were potentially to be included in the final model. I used this elimination scheme because all the risk factors in the data set are variables that could be relevant to the model, but by

convention if the predictor has a P-value greater than 0.25 in a univariate analysis it is highly unlikely that it will contribute anything to a model which includes the other risk factors [283].

8.4.6 Multivariable Analysis

Some variables (e.g. age) were included in the model as they were believed *a priori* to be potential confounders. The hazard ratio (HR)/Relative risk (RR) is the hazard rate among children at the higher risk level divided by the hazard rate among those at the lowest-risk level. Subsequent variables were included in the model using a non-automated forward stepwise regression [283]. Variables with P-values < 0.25 in the univariate analysis were selected using a forward selection procedure starting with the variable with the highest test statistic. A variable was retained in the model if any category of the variable had a P-value of < 0.05. After addition of any variable a backward elimination procedure was applied to each variable already in the model. Throughout the model building process changing effect estimates and possible variable interactions were monitored. New combination variables were created where appropriate. In relation to highly correlated variables, i.e. correlation coefficient of > 0.9, only one of the variables was included. To make certain that no effect of eliminated variables had been masked, after the maximum model was fitted, the eliminated predictors were added back in one at a time to the final model.

8.5 General Results

There were a total of 29,979 observations from 469 children. The characteristics of the study population are shown on Table 8.4. Figures 8.2 to 8.4 show the different nutritional indices by age of children progressively through follow up computed using the WHO child growth standards.

Table 8.4. Characteristics of population studied

Characteristics		*Number (all)*	*Percentage (all)*
Age of PCT *	13-20 years	70	14.93
	21-30 years	229	48.83
	31-40 years	130	27.72
	41-50 years	31	6.61
	51-63 years	9	1.92
Educational level of PCT	No schooling	199	25.37
	1-7 years of schooling	184	39.23
	8-12 years of schooling	146	31.13
	>12 years of schooling	20	4.26
Literacy of PCT	Able to read	302	64.39
	Unable to read	167	35.61
Occupational class:	Professional	50	10.66
	Skilled	166	35.39
	Non-skilled	174	37.31
	Trade	78	16.63
Receive outside financial assistance	Yes	91	19.40
	No	378	80.6
Household assets (Number that have) (One or more)	Phone	111	23.67
	Radio	317	67.59
	Motor bike	5	1.07
	Bicycle	233	49.68
	TV video	69	14.71
	Electricity	37	7.91
	Cows	92	19.62
	Goats	220	46.91
	Sheep	10	2.13
Housing	Owner occupied	362	77.19
	Rented/Other	107	22.81
House type	Block wall	181	38.59
	Mud wall	288	61.41
Cooking Fuel	Mainly firewood	329	70.15
	Mainly charcoal	110	23.45
	Both	12	2.56
	Non-biomass	18	3.84
Cooking location	Outside	36	7.68
	Same house as you sleep in	175	37.31
	Different house from where you sleep	258	55.01
Water source	Open/closed well	24	5.12
	Piped	423	90.19
	Both	22	4.69
Water Site	Owned	42	8.96
	Public	359	76.55
	Shared	68	14.50
Toilet type	No toilet	155	33.05
	Flush toilet	32	6.82

	Latrine	282	60.13
Number of children in the household	1 child	53	11.30
	2-3 children	154	32.84
	4-6 children	175	37.31
	7+ children	87	18.55
Household size	1-4 members	108	23.03
	5-9 members	242	51.60
	10+ members	119	25.37

*PCT-Primary care taker

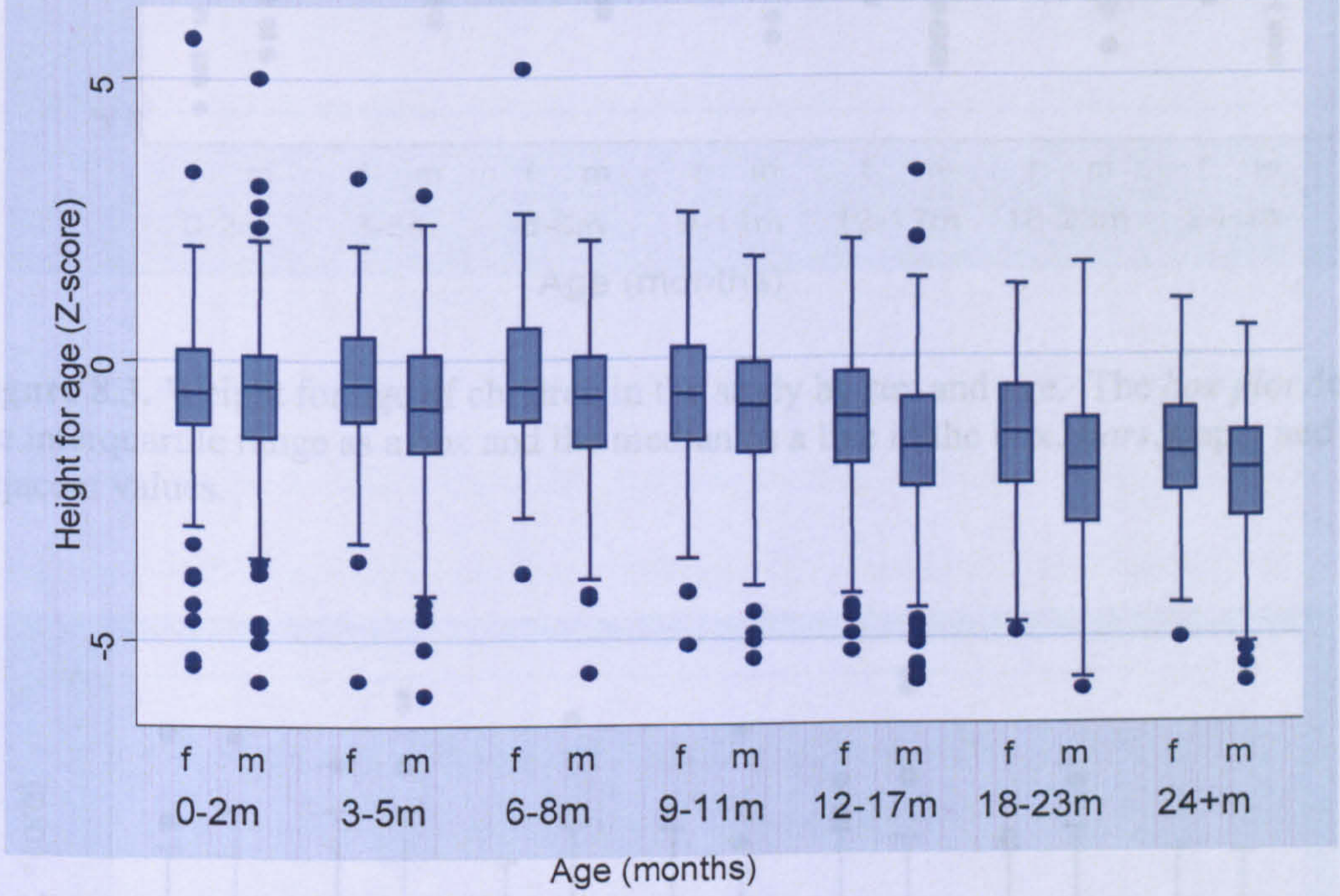


Figure 8.2. Height for age of children in the study by sex and age. The *box plot* depicts the interquartile range as a *box* and the median as a line in the box. *Bars*, upper and lower adjacent values.

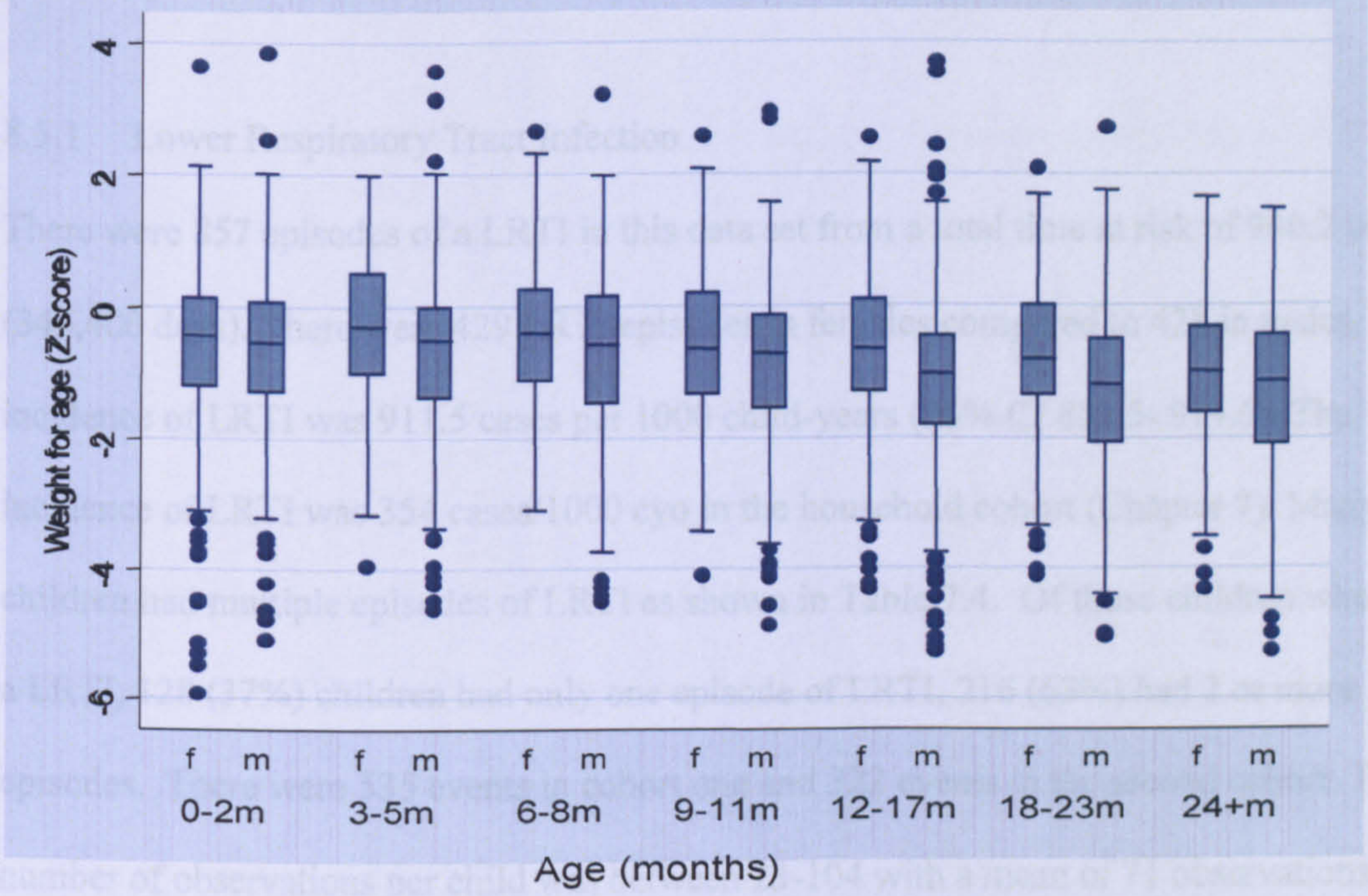


Figure 8.3. Weight for age of children in the study by sex and age. The *box plot* depicts the interquartile range as a *box* and the median as a line in the box. *Bars*, upper and lower adjacent values.

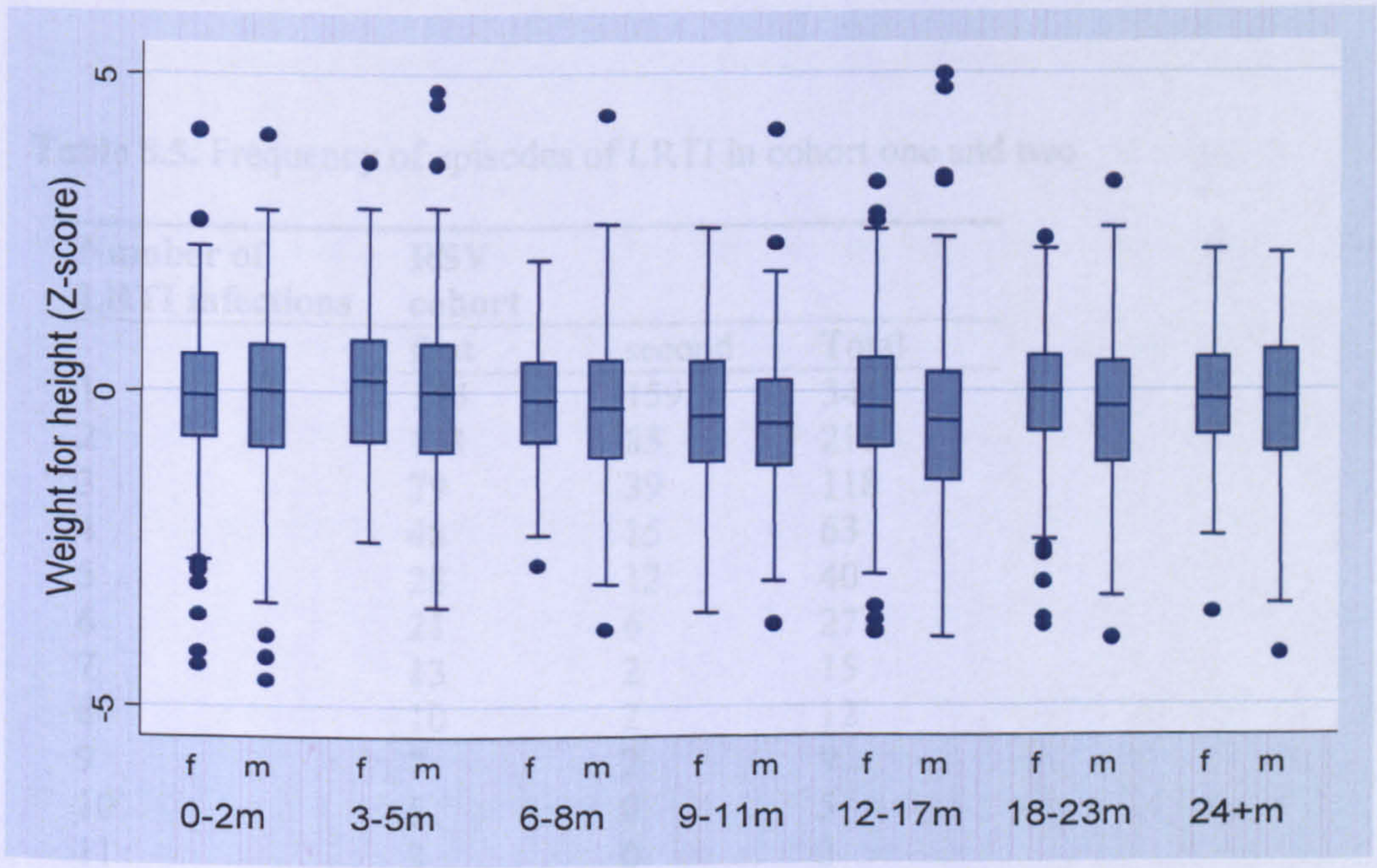


Figure 8.4. Weight for height of children in the study by sex and age. The *box plot* depicts the interquartile range as a *box* and the median as a line in the box. *Bars*, upper and lower adjacent values.

8.5.1 Lower Respiratory Tract Infection

There were 857 episodes of a LRTI in this data set from a total time at risk of 940.2 years (343,400 days). There were 429 LRTI episodes in females compared to 428 in males. The incidence of LRTI was 911.5 cases per 1000 child-years (95% CI 852.5- 974.6). The incidence of LRTI was 354 cases/1000 cyo in the household cohort (Chapter 7). Many children had multiple episodes of LRTI as shown in Table 7.4. Of those children who had a LRTI, 128 (37%) children had only one episode of LRTI, 216 (63%) had 2 or more episodes. There were 535 events in cohort one and 322 events in the second cohort. The number of observations per child was between 23-104 with a mean of 71 observations and a median of 71 observations. 22 children who had 8 or more episodes of LRTI (Table 8.5) had between 76- 101 visits i.e. higher than the average number of visits. This seems to suggest that these are children who were frequently ill and came to hospital more.

Table 8.5. Frequency of episodes of LRTI in cohort one and two

Number of LRTI infections	**RSV cohort**		
	first	second	Total
1	185	159	344
2	131	85	216
3	79	39	118
4	48	15	63
5	28	12	40
6	21	6	27
7	13	2	15
8	10	2	12
9	7	2	9
10	5	0	5
11	3	0	3
12	2	0	2
13	1	0	1
14	1	0	1
15	1	0	1
Total	535	322	857

8.5.2 RSV infection and RSV specific Lower respiratory tract infection (RSV-LRTI)

There were 362 episodes of clinical RSV in the data set from a total time at risk of 936.7 years (342,144.4 days). 283 single infections, 79 children infected more than once (68 twice, 8 three times, two four times and one five times). The incidence of clinical RSV infection was 348.6 (95% CI 349-428) cases per 1000 child-years. There were 92 episodes of RSV-LRTI in the dataset. There were 86 single episodes of RSV-LRTI, 6 children had two episodes of RSV-LRTI. The overall incidence of RSV-LRTI in this study population was 98.2 (95% CI 80-121) cases per 1000 child-years.

8.6 Discussion

8.6.1 Advantages and limitation or the study design and questionnaire

The cross-sectional survey was done within a birth cohort study. Because the survey was carried out within a longitudinal study, comprehensive data on infection history was available for all children and therefore yields more information. The survey was carried out in 2004 just before the end of follow up of cohort one. This meant that some children had been lost to follow up before the risk factor study and data on their risk factors was unavailable. Out of a total of 635 birth cohort children who had at least one visit, 469 (74%) children were part of the survey.

The survey was done at one point in time with some of the data being collected retrospectively. As some risk factors studied changed over the follow up time, an improved study design would be a prospective study collecting risk factor data at the beginning of the study and have this updated periodically through out the study period. In this way the recall bias would have been reduced. One important point to note is that during the RSV epidemics (around which the survey questions were structured; Appendix S, p 358) the

active surveillance was by weekly home visit, thus mothers were more likely to recall events of that particular time period thus potentially reducing the level of recall bias.

Even though the questionnaire had been piloted, several questions were found to be redundant at the time of analysis: sex of the primary care taker; 99% were female, owning a motorbike (1%), obtaining drinking water from an open source (river) <1%. Thus these could have been excluded from the questionnaire, although they were considered to be variable prior to analysis of results. These questions may not be considered necessary in subsequent surveys. In addition, an improved definition of the duration of breastfeeding is recommended clarifying exclusive breastfeeding from weaning and breastfeeding. Similarly, several other socio-economic questions ought to be included to improve computation of the index e.g. land ownership, type of floor and type of roofing material used, measure of crowding index. Overall we were able to capture adequate information based on the questions included in the questionnaire.

There are three outcomes for which this survey was conducted; clinical RSV infection, RSV-LRTI and all cause LRTI. Outcome was ascertained based on recommended methods. LRTI was defined according to modified WHO guidelines adjusted for low specificity particularly in children < 2 months [5]. The incidence of LRTI in this study is similar to results from other studies from developing countries; discussed previously (see Chapter 6), implying an adequate detection of cases in this study. The method of detection of RSV antigen in nasal specimens and its limitations has been discussed in Chapter 3. The IFAT test used for RSV diagnosis is the recommended diagnostic method for samples from children [242]. It is still possible that sub-clinical RSV infection may have occurred resulting in lower incidence of RSV infection and disease. Serological assays currently

underway will shed more light on this concern. Consequently, assessment of risk factors in this thesis was limited to clinical infection, Chapter 8.

8.6.2 Validity of asset index

There is a lack of clarity in the determination of how many components to retain. Usually one weighs the cost of using additional components in subsequent analysis against the benefits of the additional variance for which they account. However, in the computation of asset indices it is generally assumed that the first component captures an adequate measure of welfare of a household [284, 285]. To test this assumption and the robustness of asset index we use the rank correlation coefficient, which compares the degree to which two methods (i.e. index computed using the first component and that computed using 3 components, Table 8.6) produce the same ranking of households. The correlation between these two indices was 0.78; $p < 0.001$. Thus only one component was retained.

Table 8.6. Classification differences of the study population using different indices

3 Quantiles computed from first component	**3 quantiles computed from first 3 PCs**			
	1	2	3	Total
1	147	31	2	180
%	*77.78*	*18.9*	*1.14*	*34.03*
2	38	101	35	174
%	*20.11*	*61.59*	*19.89*	*32.89*
3	4	32	139	175
%	*2.12*	*19.51*	*78.98*	*33.08*
Total	189	164	176	529
	100	100	100	100

A possible problem with the asset index is the possibility that it may tend to reflect community variables especially locally available infrastructure rather than actual variables that are specific to households [270], or specific risks to the outcomes. It is reassuring then

to know that there is a clear difference in variables not related to infrastructure like the use of blocks in building homes or in actual household consumables like phones and radios. However, a comparison of results obtained from the asset index with income and expenditure or consumption data of each household would be the ideal means of comparing the correlation of the asset index. Thus to further validate the asset index, we compared the occupational category of the major income provider (MIP) in the three socio-groupings, Table 8.7. There was significant association (Pearson's χ^2, $P< 0.001$) between the index and the profession of the MIP. "Rich" households had a higher proportion of MIPs in the professional and skilled category. Individuals in these professions have higher earnings thus are considered to be better off. "Poor" households had a high proportion of un-skilled MIPs who were mostly manual laborers with minimal and inconsistent income. This result thus adds weight to the validity of the asset index.

Table 8.7 Comparison between the asset index and the profession of a household's major income provider.

	Poverty Classification			
Job description of MIP	**poor**	**middle**	**rich**	**Total**
non-skilled	83	52	40	175
%	*47.43*	*29.71*	*22.86*	*100*
trade	20	33	25	78
%	*25.64*	*42.31*	*32.05*	*100*
skilled	46	54	66	166
%	*27.71*	*32.53*	*39.76*	*100*
professional	1	13	36	50
%	*2*	*26*	*72*	*100*
Total	150	152	167	469
	31.98	32.41	35.61	100

MIP-major income provider

Some of the assets were found to follow a non-linear relationship (Table 8.3). These are the assets that one would say are somewhat affordable but still out of reach for the very poor households and include owning a bicycle or owning goats and sheep. Most homesteads in this community will own a few goats and/or sheep especially those that are out of the main township area. Furthermore, most of the public transport is restricted to the main highway thus owning a bicycle is convenient in terms of facilitating movement internally thus becomes almost a necessity. A bicycle is a relatively cheap mode of transport and requires only a single initial investment.

The children in this study were for the most part well-nourished. Boys had consistently lower median z-scores than girls. As is usually the case at birth and up to the age of 1 year, the children in this study were within the limits considered to be well nourished (z-score >1) [286]. Although weight-for-age (underweight) and height-for-age (stunting) scores reduced in older children in comparison with the WHO standard, this decline was moderate with very few children being classified in the severely underweight or stunting categories (≤-3). This fall was more pronounced for stunting (Figure 8.2). The decline in the weight-for-height was less pronounced. From this data we can conclude that the incidence of malnutrition in the birth cohort was generally low. However, it is worth mentioning the likely inadequacy of the current measurements. The anthropometric measurements were obtained every three months, which is considered to be inadequate to describe the rapid and changing rate of growth in early infancy [274]. Ideally, measurements should be taken at shorter intervals to capture this rapid growth.

Chapter Nine

Determinants of Respiratory Disease at Household and Individual level

9.1 Introduction

The focus of the study in these risk factors chapters is the infant i.e. birth cohort child – as a key group for prevention of disease. Variations in disease incidence [5, 6, 11, 40, 126, 127] that cannot be accounted for by methodologically differences can probably be associated to existing differences in socio-economic-cultural risk factors for infection and disease for example family size and structure, schooling, religious practices [11], environmental pollution, nutritional status, co-infection [41, 207] and host genetics [40] (detailed in sec 2.9, Chapter 2). These risk factors are of considerable intrinsic interest and some of these were explored in this population.

9.2 Chapter aims

In this chapter, the birth cohort data presented in Chapter 8 are initially used to determine the risk factors for RSV infection (RSVI). Since it is well-known that every child gets RSV in the first couple of years of life this was specifically to identify risk factors for clinical infection (ARI). Those risk factors identified for clinical RSVI were then used as a template to characterize the risk factors for RSV specific lower respiratory tract infection (RSV-LRTI). The identified risk factors for RSV-LRTI are then compared with those determined for all-cause LRTI.

9.3 Age-stratified incidence estimates

Age-stratified incidence of RSVI and disease (RSV-LRTI)

The age-stratified incidence of RSVI and disease (RSV-LRTI and severe RSV-LRTI) is shown in Figure 9.1. There were no observed differences in RSV infection and RSV-LRTI

rates by age. The incidence of severe RSV-LRTI appears to start declining from the age of 9 months onwards, reaching a significant decline at 12-17 months and beyond (Wald test, P-value < 0.05).

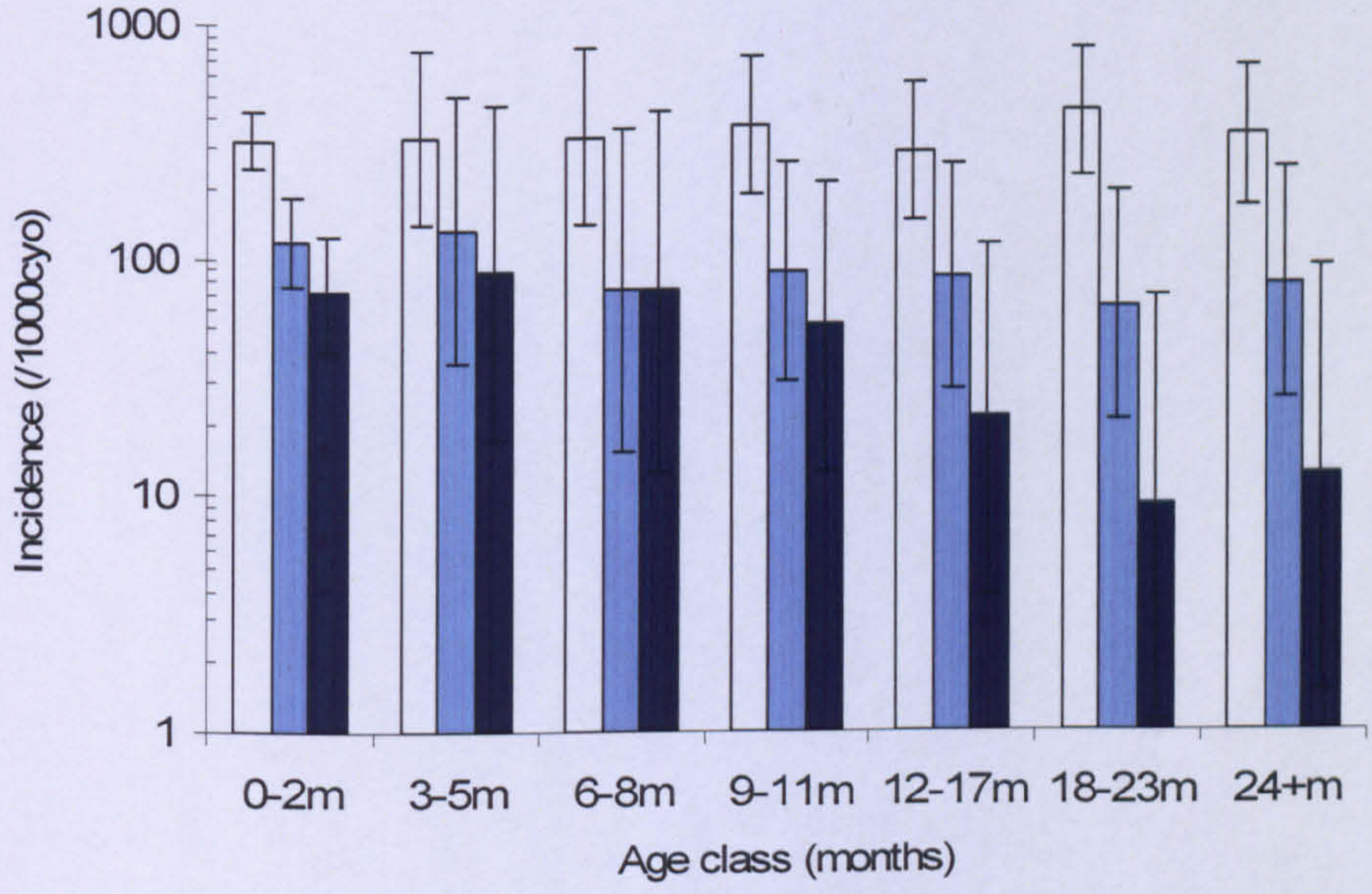


Figure 9.1 Age-stratified incidence of RSV infection (open bars), RSV-LRTI (light blue bars) and severe RSV-LRTI (dark blue bars).

Age-stratified incidence of all-cause LRTI

Figure 9.2 shows the age stratified incidence of all-cause LRTI cases. The incidence of LRTI did not significantly differ by age group. The peak incidence was in children less than 6 months and 12-17 months of age.

Figure 9.2. Incidence of all-cause LRTI by age class with 95% CI

9.3.1 Interactions between age and gender, poverty level and HH density

Interactions between age and gender, age and poverty classification and age and household density were investigated for all three outcomes; RSVI, RSV-LRTI and all-cause LRTI. No differences in infection and disease rates were observed between boys and girls for all outcomes. The incidence by age and poverty classification for the three outcomes together is shown in Figure 9.3. Slightly higher rates of disease were observed in certain age groups in poorer households. The incidence of RSVI and LRTI was generally the same in the different households (Wald test, $P>0.05$). The incidence was lower (borderline significance) in children 12-17 and 18+ months old (Wald test, $P=0.089$ and 0.065) respectively in households classified as rich.

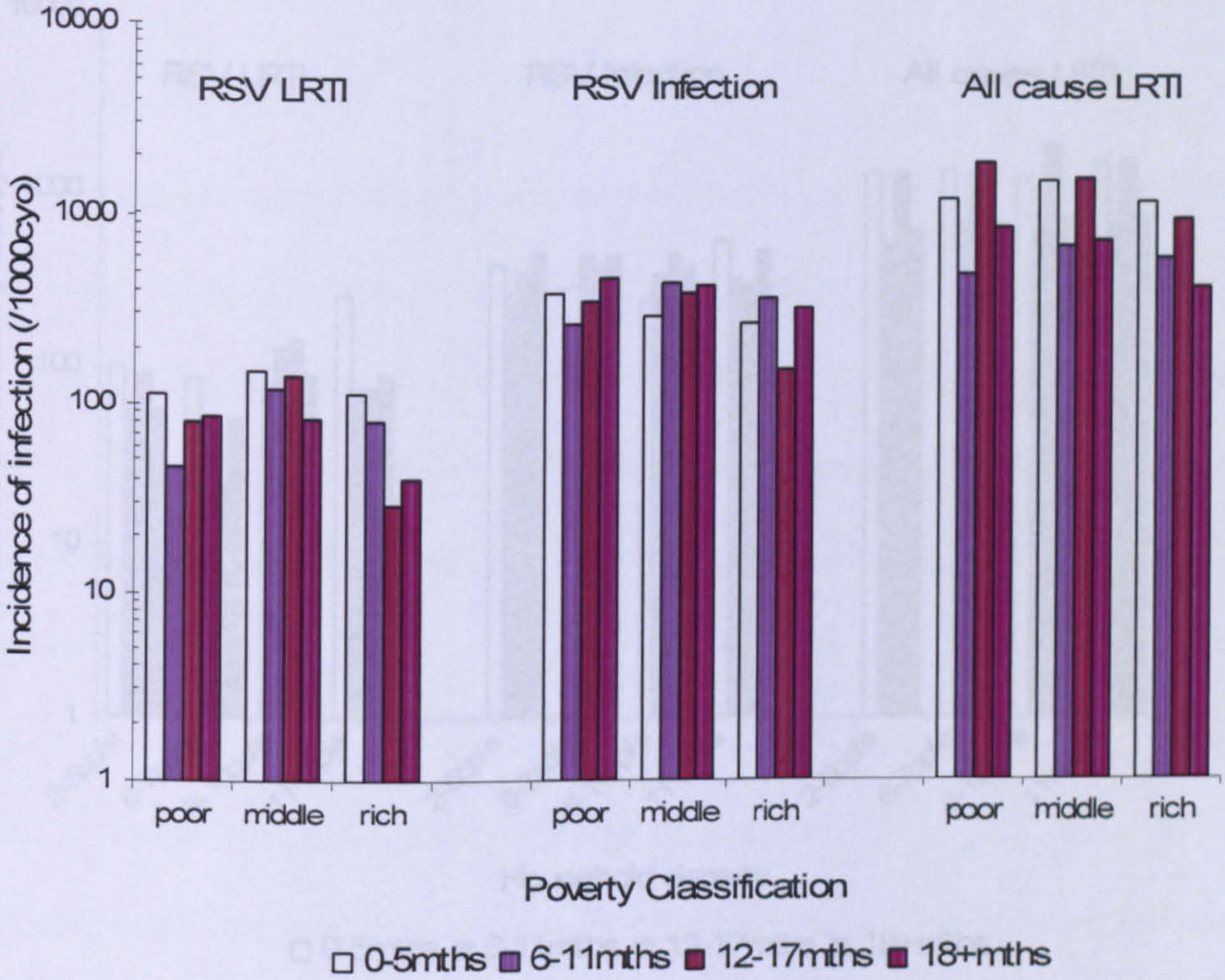


Figure 9.3 Age-stratified incidences of RSV-LRTI, RSVI and LRTI cases by poverty classification: poor, middle and rich.

Figure 9.4 shows the incidence rates by age class for households of different size. There were no significant differences observed in the incidence rates for RSVI and LRTI in the different ages in the different households. However, there was a suggestion of decreased rates of RSV-LRTI in older children: 6-11 months (P=0.035), 12-17 months (P=0.063) and 18+ months (P=0.029) in larger households (11+ people).

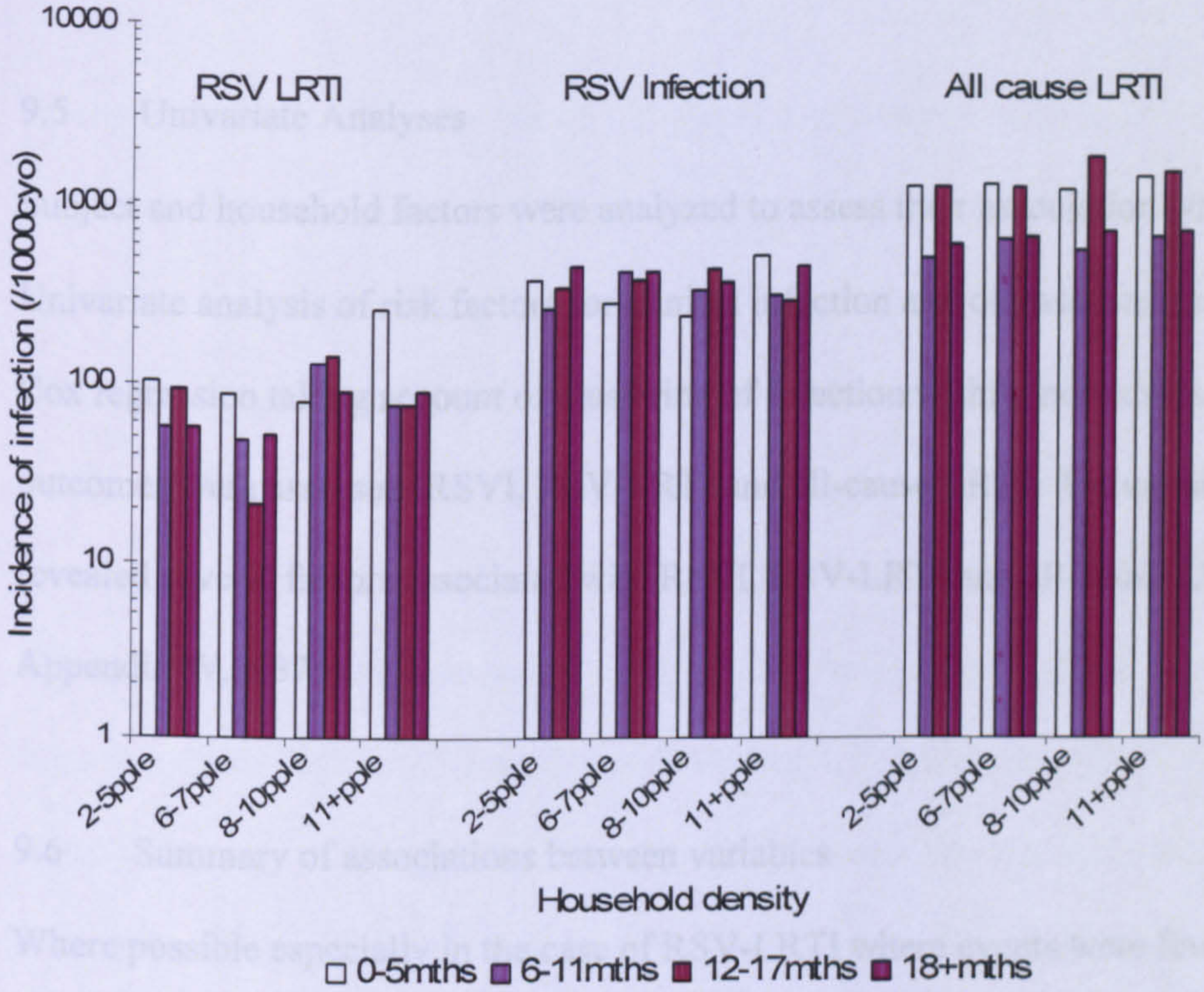


Figure 9.4. Age-stratified incidence of RSV-LRTI, RSVI and LRTI by household density

9.4 Exploring data for RSVI, RSV-LRTI and all-cause LRTI outcomes

Screening was done using a log rank test to test for equality of survival function across strata (described in Chapter 8) for all the outcome variables. Those variables that gave a p value <0.25 (conventional cutoff) had potential to be included in the final model. Tobacco exposure was defined as being present if any person living in the household smoked. Breast-feeding was defined as being present if the subject was fed by breast at all.

9.5 Univariate Analyses

Subject and household factors were analyzed to assess their association with disease. Univariate analysis of risk factors for clinical infection and disease was performed using Cox regression taking account of clustering of infection within individuals. Three different outcomes were assessed: RSVI, RSV-LRTI and all-cause LRTI. Univariate analysis revealed several factors associated with RSVI, RSV-LRTI and all-cause LRTI (see Appendix W, p 375).

9.6 Summary of associations between variables

Where possible especially in the case of RSV-LRTI where events were few, categories with small numbers of observations were combined or redefined to achieve stable estimates in the multivariate model. This included the PCT age classification where the baseline age class was changed from 13-20 to 13-22 and the two higher age classes were combined creating one class of PCTs aged 41-63. The two lower categories defining birth weight were also combined to have <2.5 kg as the baseline. The baseline category for main fuel used for cooking was changed from non-biomass to charcoal as there were no RSV-LRTI events in the non-biomass category. Ownership of water site was also changed from three categories to having only two categories; owned/shared water source and public water source. For some variables this was not possible as the categories within the variables define different effects (e.g. that of care giver) and thus were left unchanged.

Certain variables were found to be associated with each other some of which were found to be highly correlated creating problems of collinearity which caused difficulties in distinguishing individual effects. By their definition variables illustrating contact intensity within the household such as the number of siblings sleeping in the same bed were

inherently collinear. The number of siblings of a certain age sleeping in the same bed was conditioned by the number of siblings of that age sleeping in the same room (variable denoting numbers sleeping in bed could only be equal to or less than number of siblings sleeping in the same room). In terms of transmission one would expect that closer contact would make transmission more likely to take place. Possible differences in risk between sleeping in same bed and/or in room with siblings were investigated. For these variables; siblings 6+ years sleeping in same house, room or bed and siblings <6 years sleeping in same house, room or bed, it was decided to include, after fitting each separately, the variable that was most significant and use this to capture the effect of the other variables. Over 50% of the siblings who are 6 years of age or older, though sleeping in the same house are sleeping in a different room from the cohort child which precludes sleeping in the same bed.

Certain socioeconomic variables were found to be associated with each other; type of cooking fuel, house ownership, water site location and ownership, type of house and type of toilet. A new variable combining the type of house and sanitation was created. Majority of owned houses are mud walled. Ordinarily owning a home would imply higher socio-economic status but in this environment this usually refers to the local rural homestead thus in a majority of the situations there is no financial investment associated with ownership evidenced by the fact that 41% of the owner occupied homes were classified as "poor". Those that are rented are mostly block walled and are households that are better off; 91% classified as "rich". Moreover, interaction between the composite variable denoting poverty class and those variables used in its computation was observed. For these, only those that were associated with infection or disease in the univariate Cox model were assessed for inclusion into the final model. Each of these socioeconomic predictors was first assessed in

the multivariate model separately and then in combination with the other variables and those still remaining significant were included in the final model

There was a highly significant correlation between number of children in the household and family size (r=0.9), thus only one of these was included in the multivariate model as was the case between literacy and number of years of education (r=0.8) of the primary care taker (PCT). We chose to keep family children and education as each one automatically explains the other and these remained significant in the multivariate model. We evaluated possible interaction between family children, family size and poverty classification, Figure 9.5. There does not appear to be an interaction just potential for confounding i.e. the nature of the association (slope) of family children and family size does not seem to alter for different poverty class, the linear relationship for each poverty status has similar slope and the lines do not cross over.

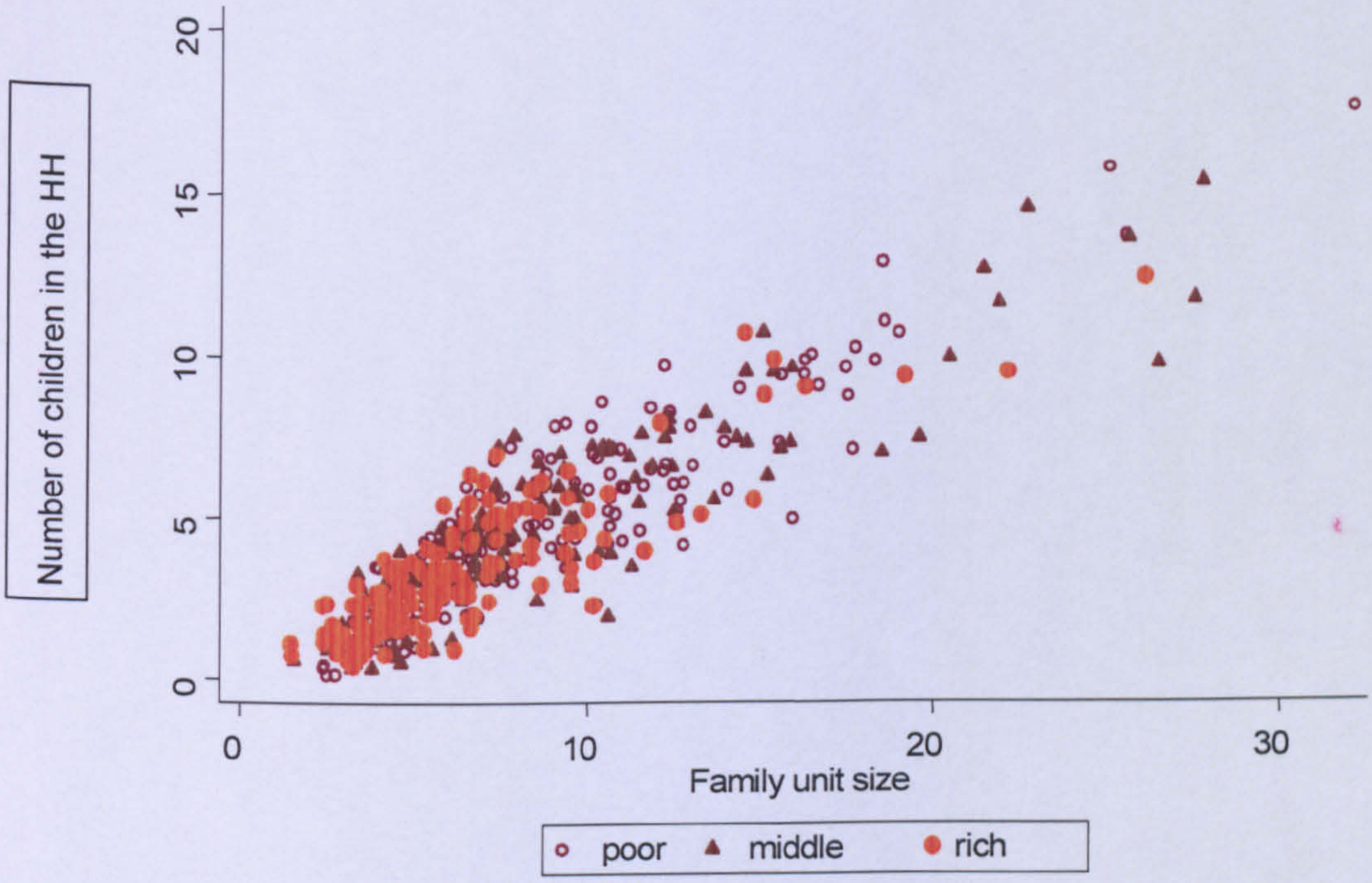


Figure 9.5. Correlation between family density and number of children in the household (y-axis) for households of difference socioeconomic class.

9.7 Analyses of Independent Effects

Independent predictors of disease were analyzed using multivariate models using a forward stepwise approach. Out of 42 variables 12, 21, 33 had a significant log-rank test of equality across strata for clinical RSVI, RSV-LRTI and LRTI respectively, Table 9.1 and 9.2. These variables were included as potential candidates for the final model for the three outcome measures respectively some being excluded as discussed above. In this analysis we used forward stepwise regression to build the maximum model always keeping in mind biologically plausible relationships. A shared frailty model was used for all-cause LRTI outcome with significant within-child correlation. This was not the case for RSVI or RSV-LRTI for which robust standard errors adjusted for clustering of measurements within-child were calculated. The risk factors that independently predicted increased risk of disease are shown and discussed below. In each analysis, P-values are for the tests of the null

hypothesis that each hazard ratio is one. Significant P-values are in bold whereas those of borderline effects are underlined.

Table 9.1. Estimated hazard ratios for factors associated with clinical RSV infection.

Risk Factor	*Categories (events)*	*cyo*	*Hazard Ratio (95% CI)*	*P>z*
PCT age group _0	13-20 years (49)	137	-	-
PCT age group _1	21-30 years (186)	465	1.15 (0.87, 1.52)	0.316
PCT age group _2	31-40 years (92)	263	1.09 (0.79, 1.49)	0.600
PCT age group _3	41-50 years (23)	63	1.02 (0.66, 1.57)	0.925
PCT age group _4	**51-63 years (12)**	19	**2.13 (1.27, 3.59)**	**0.004**
care giver_0	Mother (300)	767	-	-
care giver _1	other family member (41)	115	0.77 (0.54, 1.08)	0.130
care giver _2	**House help (7)**	22	**0.54 (0.30, 0.99)**	**0.047**
care giver _3	School (1)	2	1.06 (0.87, 1.30)	0.571
care giver _ 4	mother / other family member(13)	38	1.28 (0.80, 2.05)	0.312
smokers in HH_0	None (255)	657	-	-
smokers in HH_1	1 (83)	222	0.92 (0.74, 1.15)	0.477
smokers in HH_2	**2 or more (24)**	44	**1.40 (1.07, 1.84)**	**0.016**
family assisted	No (273)	767	-	-
family assisted	Yes (89)	183	1.20 (0.98, 1.46)	<u>0.084</u>
house_toilet_0	mudwall no toilet (124)	274	-	-
house_toilet_1	blockwall no toilet (15)	30	1.03 (0.73, 1.46)	0.870
house_toilet_2	mudwall latrine (121)	301	0.98 (0.79, 1.21)	0.826
house_toilet_3	blockwall latrine (88)	268	0.83 (0.65, 1.06)	0.141
house_toilet_4	mudwall flush toilet (1)	5	0.53 (0.19, 1.47)	0.221
house_toilet_5	**block flush toilet (13)**	57	**0.60 (0.40, 0.91)**	**0.016**

Notes (apply to all subsequent tables)
Underlined values mean borderline significance
Bold co-variates indicate significant results

Increased age of the PCT (>50) was the strongest predictor of increased risk of RSVI, also thought to be related to household size. Other factors not related to PCT but associated with increased risk of RSVI were exposure to tobacco smoke and two potential indicators of higher socio-economic status (SES): house type and toilet (block walled house with a flush toilet) and the birth cohort child's care giver (hired house help). Other SES indicators that were significant in the univariate e.g. profession of the major income provide (MIP) and the variables representing crowding and intensity of contact (Appendix W, p 375) were not independently correlated to RSVI. The significant variables i.e. those that predict the risk of infection were included in the RSV-LRTI multivariate analysis as the baseline. Additional variables were then fitted in an attempt to explain the difference between infection and disease. Results from this multivariate regression are shown in Table 9.2.

Table 9.2. Risk factors independently predicting increased hazard of RSV-LRTI / disease

Risk Factor	*Categories*	*cyo*	*Hazard Ratio (95% CI)*	*P>z*
child's care_0	mother (73)	767	-	-
child's care _1	another family member (13)	115	0.95 (0.53, 1.71)	0.872
child's care _2	house help (3)	22	1.49 (0.56, 4.00)	0.424
child's care _3	school (0)	2	0.00 (-)	.
child's care _ 4	mother / family member (3)	38	1.26 (0.47, 3.37)	0.646
PCT*_agegroup_0	13-24 years (26)	329	-	-
PCT_agegroup_1	25-30 years (24)	274	1.03 (0.58,1.82)	0.928
PCT_agegroup_2	31-40 years (28)	263	1.49 (0.88, 2.53)	0.137
PCT_agegroup_3	**41-63 years (14)**	82	**2.19 (1.13, 4.25)**	**0.021**
smokers in HH_0	None (66)	657	-	-
smokers in HH_1	1 (21)	222	0.87 (0.53, 1.45)	0.597
smokers in HH_2	2 or more (5)	44	0.81 (0.36, 1.84)	0.621
height for age z-score _0	>-1 (38)	465	-	-
height for age z-score _1	-1.99 to -1 (26)	246	1.34 (0.83, 2.17)	0.227
Height for age z-score_2	**≤ -2 (28)**	233	**1.73 (1.08, 2.76)**	**0.022**
family assisted	No (69)	767	-	-
family assisted	Yes (23)	183	1.34 (0.86, 2.09)	0.195
house_toilet_0	mudwall no toilet (26)	274	-	-
house_toilet_1	**blockwall no toilet (7)**	30	**3.62**(1.53, 8.88)	**0.003**
house_toilet_2	mudwall latrine (34)	301	1.57 (0.96, 2.57)	0.074
house_toilet_3	blockwall latrine (24)	268	1.41 (0.79, 2.53)	0.248
house_toilet_4	mudwall flush toilet (0)	5	0.00 (-)	.
house_toilet_5	block flush toilet (1)	57	0.25 (0.03, 2.11)	0.203
family_children_0	1-5 children (60)	684	-	-
family_children_1	6-10 children (26)	230	0.97 (0.57, 1.66)	0.920
family_children_2	**11+ children (6)**	27	**2.58 (1.03, 6.50)**	**0.044**
siblings under 6 years_0	no siblings< 6 yrs (16)	263	-	-
siblings under 6 years_1	**1-2 siblings<6 yrs (58)**	548	**2.00 (1.17, 3.42)**	**0.011**
siblings under 6 years_2	3-4 siblings<6 yrs (15)	120	1.99 (0.81, 4.91)	0.133
siblings under 6	5+ siblings<6 yrs (3)	21	1.74 (0.54, 5.63)	0.356

years_3				
job description of MIP_0	non-skilled (35)	356	-	-
job description of MIP_1	Trade (21)	156	1.64 (0.98, 2.73)	0.058
job description of MIP_2	Skilled (27)	329	0.91 (0.56, 2.09)	0.701
job description of MIP_3	Professional (9)	99	1.04 (0.48, 2.27)	0.916

Of the predictors of RSVI, increased age of PCT (>40 years) and house and toilet type were also correlated with an increase in risk of RSV disease. Of the traditional risk factors for RSV disease - crowding (number of children in the home) and characteristics of contact person (number of children < 6 years in the home) were found to correlate with increased risk of RSV-LRTI. Stunting (height-for-age z-score (haz ≤ -2) was a significant independent predictor of RSV disease. Other factors found to be associated with increased risk of disease in the univariate model (Appendix W, p 379) were not confirmed as independent predictors of RSV-LRTI indicating their probable association with other risk factors (e.g. number of children in the home and house/toilet type), house ownership, profession of MIP (both SES indictors) and other variables also related to crowding; number of siblings of different ages were not found to be independent predictors of increased risk of RSV-LRTI. In an attempt to compare differences in the risk factors for RSV-LRTI and those for all-cause LRTI we sought to determine predictors of LRTI in this study population, Table 9.3.

Table 9.3. Risk factors independently predicting increased hazard of all-cause LRTI

Potential risk factors	*Categories (number of events)*	*cyo*	*Hazard Ratio (95% CI)*	*P>|z|**
current age_0	0-5 months (245)	200	-	-
current age_1	**6-11 months (120)**	214	**0.55 (0.41, 0.74)**	**0.000**
current age_2	**12- 17 months (290)**	211	**1.72 (1.31, 2.25)**	**0.000**
current age_3	18+ months (202)	329	0.74 (0.52, 1.04)	0.082
multiple babies_0	1 child (790)	876	-	-
multiple babies_1	Twins (50)	55	0.80 (0.54, 1.19)	0.275
multiple babies_2	**Triplets (17)**	5	**4.12 (1.55, 10.99)**	**0.005**
education_0	no schooling (249)	241	-	-
education _1	1-7 years (353)	383	0.95 (0.76, 1.19)	0.675
education _2	8-12 years (241)	301	0.81 (0.64, 1.19)	0.092
education _3	**>12 years (14)**	38	**0.40 (0.21, 0.76)**	**0.005**
Height-age-z score_0	>-1 (363)	465	-	-
height-age-z score_1	**-1.99 to -1 (239)**	246	**1.27 (1.06, 1.52)**	**0.010**
height-age-z score_2	**≤ -2 (255)**	235	**1.38 (1.12, 1.70)**	**0.002**
family_children_0	1-4 children (598)	684	-	-
family_children_1	6-10 children (214)	230	0.96 (0.77, 1.19)	0.696
family_children_2	**11+children (45)**	27	**1.68 (1.07, 2.63)**	**0.023**
main_fuel_0	gas/paraffin (16)	36	-	-
main_fuel_1	Charcoal (158)	211	1.34 (0.74, 2.44)	0.339
main_fuel_2	Firewood (660)	657	1.71 (0.96, 3.05	0.071
main_fuel_3	firewood/charcoal(23)	23	2.11 (0.97, 4.60)	0.060
# siblings <6yr sleeping in same room as index_0	no siblings < 6 yrs (411)	520	-	-
# siblings <6yr sleeping in same room as index_1	1 -2 sibling < 6 yrs (408)	411	1.19 (0.99, 1.43)	0.067
# siblings <6yr sleeping in same room as index_2	3 siblings < 6 yrs (11)	11	1.69 (0.81, 3.51)	0.162
# siblings 6+ years sleeping in same room as index_0	no siblings 6+ years (535)	630	-	-
# siblings 6+ years sleeping in same room as index_1	**1 siblings (194)**	183	**1.29 (1.04, 1.61)**	**0.021**

# siblings 6+ years sleeping in same room as index_2	2-3 siblings (88)	110	0.83 (0.62, 1.11)	0.213
# siblings 6+ years sleeping in same room as index_3	4 siblings (13)	18	0.63 (0.31, 1.28)	0.205
# siblings <6yr going to school_0	no siblings < 6 yrs going to school (709)	794	-	-
# siblings <6yr going to school_1	1 sibling < 6 yrs going to school (86)	131	0.80 (0.62, 1.05)	0.107
# siblings <6yr going to school_2	2-3 siblings < 6 yrs going to school (35)	27	1.52 (0.98, 2.33)	0.059

Age was an independent predictor of LRTI. The effect of age was mixed, associated with protection in those 6-11 months and increased risk in those 12- 17 months. Being a child of a multiple birth was the strongest independent predictor of LRTI probably reflecting its association with lower birth weight which was not found to be associated with LRTI in this study. Other factors found to be associated with increased risk of LRTI were height-for-age z-score ≤-1, crowding (number of children in the home), intensity of contact with siblings < 6 years of age (number sleeping the same room as cohort child). Borderline significant factor was having 2-3 siblings <6 years attending school. Two factors were found to be associated with protection from LRTI, having a care taker with a college education (>12 years of schooling) and having a young sibling (<6 years) attending school. Reflecting the possible association of these variables with each other and with other variables already in the model, certain variables that were significant at univariate analyses did not make it into the multivariate model (Appendix W, p 375). These include weight-for-height z-score (related to height for age z-score), house and toilet type, the cohort child's care giver (related to SES), number of siblings < 6 years and number of male siblings (related number of children in the home). These variables are related and probably ultimately related to socioeconomic status. Because of the small number of multiple births

in these data, the LRTI multiple regression model was re-fitted excluding this variable to achieve more stable estimates. There was minimal difference in the effect estimate, results not shown. A summary table comparing individual risk factors across the different outcome groups is provided below (Table 9.4).

Figure 9.4 Summary table comparing individual risk factors across the different outcome groups is provided below

Risk Factor	*Categories*	*RSV-LRTI HR (95% CI)*	*RSV-Infection HR (95% CI)*	*All cause -LRTI HR (95% CI)*
Current age	0-5 months	-	-	-
	6-11 months	-	-	**0.55 (0.41, 0.74)**
	12- 17 months	-	-	**1.72 (1.31, 2.25)**
	18+ months	-	-	0.74 (0.52, 1.04)
Multiple babies	1 child	-	-	-
	Twins	-	-	0.80 (0.54, 1.19)
	Triplets	-	-	**4.12 (1.55, 10.99)**
Education level of PCT	no schooling	-	-	-
	1-7 years	-	-	0.95 (0.76, 1.19)
	8-12 years	-	-	0.81 (0.64, 1.19)
	>12 years	-	-	**0.40 (0.21, 0.76)**
child's care	mother	-	-	-
	another family member	0.95 (0.53, 1.71)	0.77 (0.54, 1.08)	-
	house help	1.49 (0.56, 4.00)	**0.54 (0.30, 0.99)**	-
	school	0.00 (-)	1.06 (0.87, 1.30)	-
	mother / family member	1.26 (0.47, 3.37)	1.28 (0.80, 2.05)	-
PCT* age group	13-24 years	-	-	-
	25-30 years	1.03 (0.58,1.82)	1.15 (0.87, 1.52)	-
	31-40 years	1.49 (0.88, 2.53)	1.09 (0.79, 1.49)	-
	41-50 years	**2.19 (1.13, 4.25)**[&]	1.02 (0.66, 1.57)	-
	51-63 years		**2.13 (1.27, 3.59)**	-
Number of smokers in HH	None	-	-	-
	1	0.87 (0.53, 1.45)	0.92 (0.74, 1.15)	-
	2 or more	0.81 (0.36, 1.84)	**1.40 (1.07, 1.84)**	-
height for age z-score	>-1	-	-	-
	-1.99 to -1	1.34 (0.83, 2.17)	-	**1.27 (1.06, 1.52)**
	≤ -2	**1.73 (1.08, 2.76)**	-	**1.38 (1.12, 1.70)**
family assisted	No	-	-	-
family assisted	Yes	1.34 (0.86, 2.09)	1.20 (0.98, 1.46)	-
House and toilet type	mud wall no toilet	-	-	-
	block wall no toilet	**3.62**(1.53, 8.88)	1.03 (0.73, 1.46)	-
	mud wall latrine	1.57 (0.96, 2.57)	0.98 (0.79, 1.21)	-

	block wall latrine	1.41 (0.79, 2.53)	0.83 (0.65, 1.06)	-
	mud wall flush toilet	0.00 (-)	0.53 (0.19, 1.47)	-
	block flush toilet	0.25 (0.03, 2.11)	**0.60 (0.40, 0.91)**	-
main fuel used for cooking	gas/paraffin	-	-	
	Charcoal	-	-	1.34 (0.74, 2.44)
	Firewood	-	-	1.71 (0.96, 3.05
	firewood/charcoal	-	-	2.11 (0.97, 4.60)
job description of MIP	non-skilled	-	-	-
	Trade	1.64 (0.98, 2.73)	-	-
	Skilled	0.91 (0.56, 2.09)	-	-
	Professional	1.04 (0.48, 2.27)	-	-
Number of family children	1-5 children	-	-	-
	6-10 children	0.97 (0.57, 1.66)	-	0.96 (0.77, 1.19)
	11+ children	**2.58 (1.03, 6.50)**	-	**1.68 (1.07, 2.63)**
Number of siblings under 6 years	no siblings< 6 yrs	-	-	-
	1-2 siblings<6 yrs	**2.00 (1.17, 3.42)**	-	-
	3-4 siblings<6 yrs	1.99 (0.81, 4.91)	-	-
	5+ siblings<6 yrs	1.74 (0.54, 5.63)	-	-
# siblings <6yr sleeping in same room as index	no siblings < 6 yrs	-	-	-
	1 -2 sibling < 6 yrs	-	-	1.19 (0.99, 1.43)
	3 siblings < 6 yrs	-	-	1.69 (0.81, 3.51)
# siblings 6+ years sleeping in same room as index	no siblings 6+ years	-	-	-
	1 siblings	-	-	**1.29 (1.04, 1.61)**
	2-3 siblings	-	-	0.83 (0.62, 1.11)
	4 siblings	-	-	0.63 (0.31, 1.28)
# siblings <6yr going to school	no siblings < 6 yrs going to school	-	-	-
	1 sibling < 6 yrs going to school	-	-	0.80 (0.62, 1.05)
	2-3 siblings < 6 yrs going to school	-	-	1.52 (0.98, 2.33)

9.8 Testing the proportional hazard assumption (PHA)

The fundamental assumption of the Cox regression model is that the effects of the parameters included in the model do not change during the time covered by the study. The Schoenfeld and scaled Schoenfled residuals were used to test that the model satisfied the proportionality assumption. This involves testing for a non-zero slope in a generalized linear regression of the scaled Schoenfeld residuals on functions of log time. A non-zero slope is an indication of a violation of the proportional hazard assumption. Results are shown in Appendix X (p 389). The proportionality of the model as a whole as well as for each predictor was tested. If the individual predictor tests are not significant (P->0.05) we cannot reject proportionality and we assume that the model does not violate the proportional hazard assumption.

Although the global PHA test was non-significant, 3 predictors in the LRTI model had a p-value of less than 0.05. These results were cross-checked by obtaining a graph of the regression, Figure 9.6 and 9.7. A horizontal line in the graph of the residuals is an indication that we do not have a violation of the proportional assumption. Both lines are horizontal so predictors were left in the model. Thus the PHA was held for all multivariate models used.

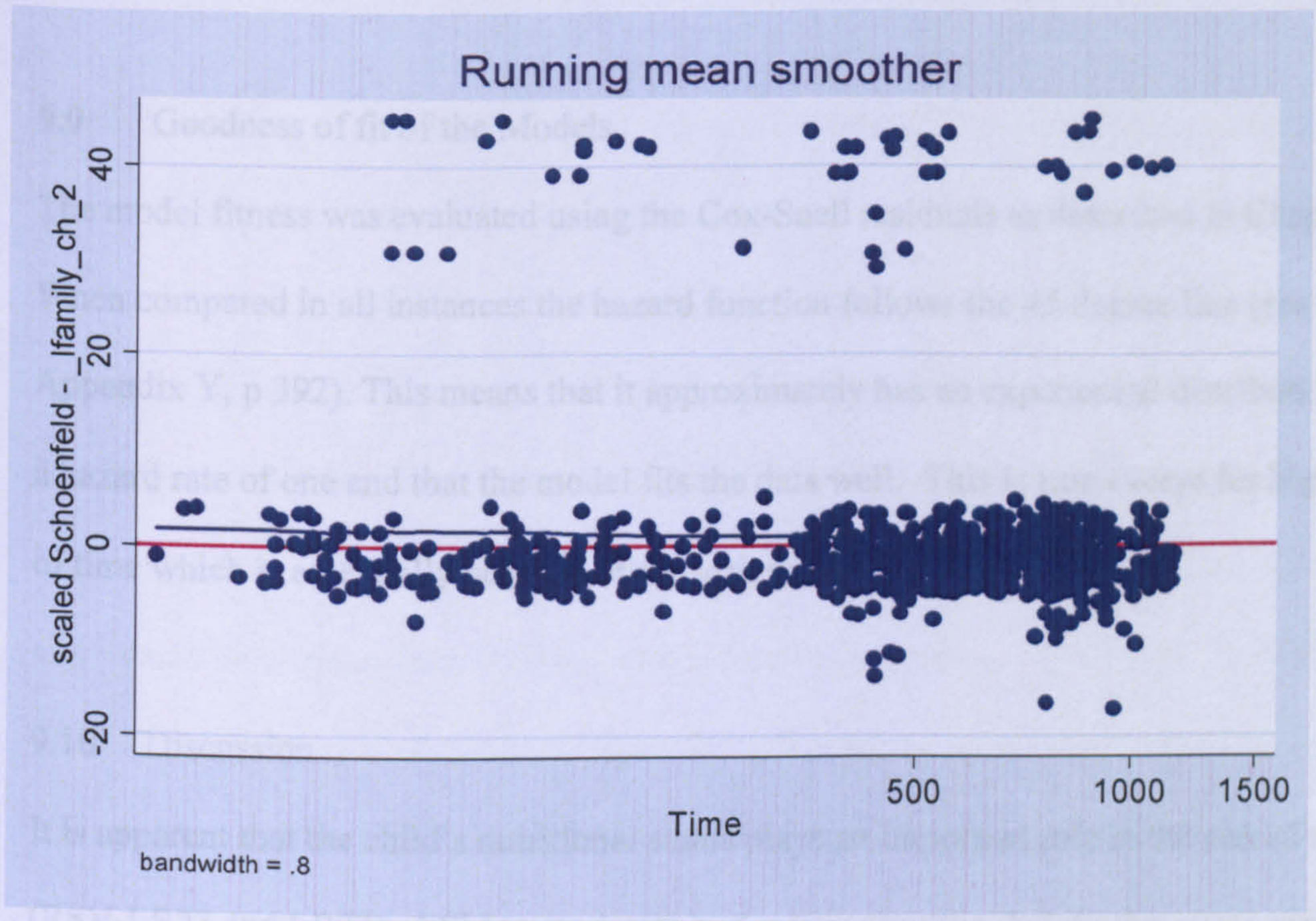


Figure 9.6 Plot for predictor number of children in the home including the reference line at y=0.

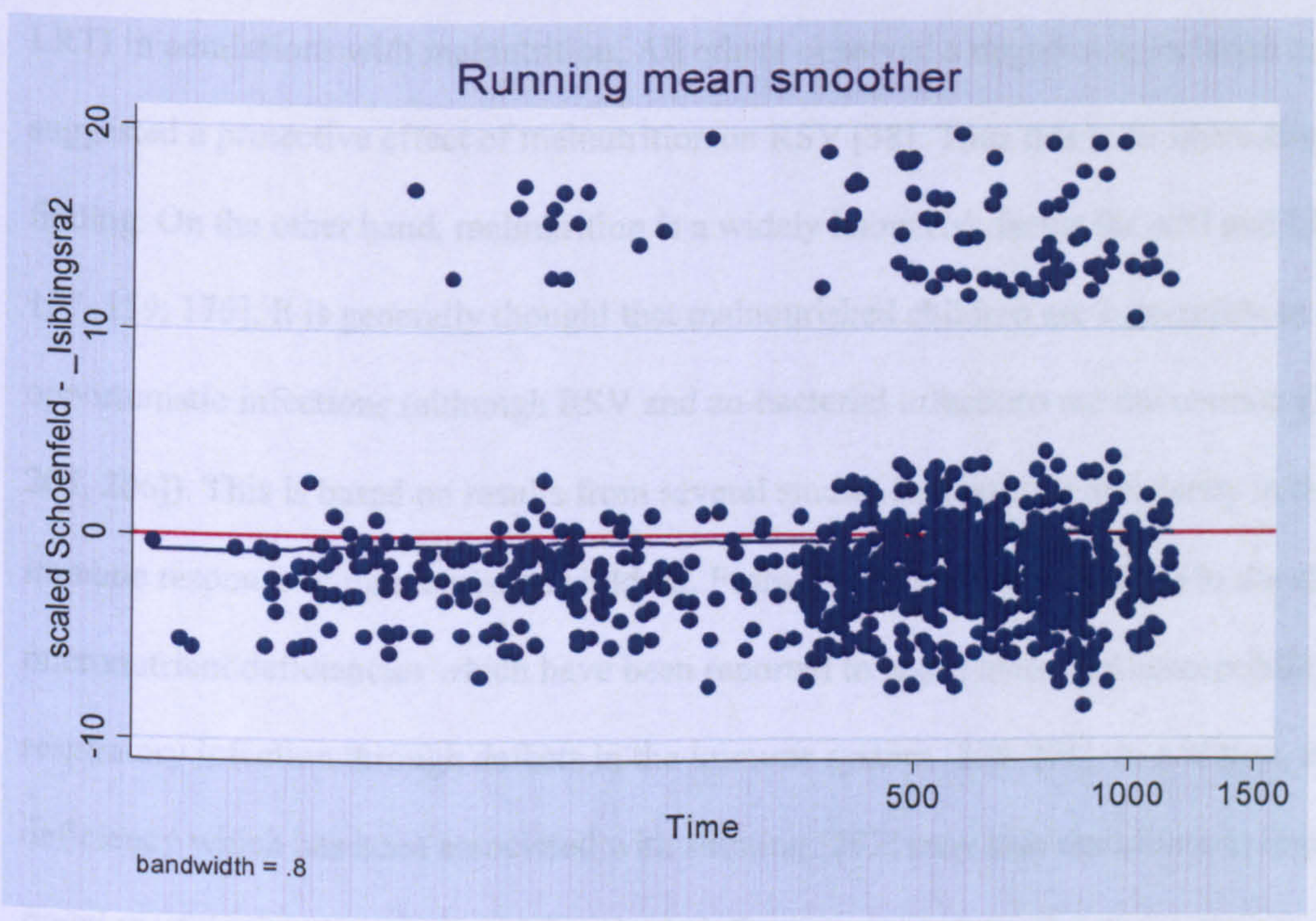


Figure 9.7 Plot for predictor number of siblings 6+ years old sleeping in same room as index including the reference line at y=0.

9.9 Goodness of fit of the Models.

The model fitness was evaluated using the Cox-Snell residuals as described in Chapter 8. When compared in all instances the hazard function follows the 45 degree line (see Appendix Y, p 392). This means that it approximately has an exponential distribution with a hazard rate of one and that the model fits the data well. This is true except for high values of time which is essentially because of censoring.

9.10 Discussion

It is apparent that the child's nutritional status plays an important role in the risk of disease (RSV-LRTI and LRTI). Mild to moderate and severe stunting (height-for-age z-score ≤ -1) was associated with increased risk of disease. This finding is consistent with only one other study in South Africa reviewed by Vardas *et al* [287] that reported increased risk of RSV-LRTI in admissions with malnutrition. All others observed a negative association or suggested a protective effect of malnutrition on RSV [38]. Thus this is an interesting finding. On the other hand, malnutrition is a widely know risk factor for ARI and LRTI [27, 157, 159, 176]. It is generally thought that malnourished children are susceptible to opportunistic infections (although RSV and co-bacterial infections are uncommon [28, 116, 205, 206]). This is based on results from several studies indicating a deficiency in the immune response in malnourished children. From T- cell dysfunction [288] to the effects of micronutrient deficiencies which have been reported to cause increased susceptibility to respiratory infection through defects in the immune system [289-291]. In addition, zinc deficiency which has been associated with stunting [292] may also contribute to impaired function of the immune system. Vitamin A deficiency caused by prolonged dietary deprivation is also said to decrease the functional integrity of the mucociliary system lining the respiratory tract [157]. Children in Kenya as a national policy routinely get vitamin A

supplements free of charge. Still, not every mother brings their children to the MCH clinic however children in this study were routinely remind about their vaccination appointments (implemented a policy of checking road to health cards at each visit). Stunting as an indication of long term malnutrition seems to be a more significant risk factor for all cause LRTI than acute (short term) malnutrition (wasting) in this study and in an earlier study done on Kenyan children [157]. Several studies have reported no association or negative association between malnutrition and incidence of RSV disease [8, 200, 202, 205, 208, 209]. The difference in results observed may be due to the inconsistency of the definition of malnutrition used and different reference standards used in the calculation of the malnutrition indices. Until recently the NCHS/WHO international reference used were based solely on US children who as infants were predominantly formula fed thus not comparable. The reference standard used in this study is from a recent WHO multi-centre study initiative using a reference computed from children in both developing and developed countries [274] and is thought to be more applicable in this setting. Similarly, the different indicators and cut offs used in studies (weight-for-age, oedema, weight-for-height, mid upper arm circumference) make direct comparison difficult. Stunting is also associated with *Trichuris* infection. Previous data from Kilifi [293] show the population to be highly parasitized with documented incidence of *Trichuris* infection of between 15-20% in children <5 years. However, the incidence of *Trichuris* infection in the study participants is unknown thus an assessment of this interaction was not possible.

To further explore this relationship, we investigated the association between the three malnutrition indicators and SES, Figure 9.8. There was no evidence of an association between malnutrition and SES although there appeared to be a suggestion that "richer" households had a higher median height-for-age z-score.

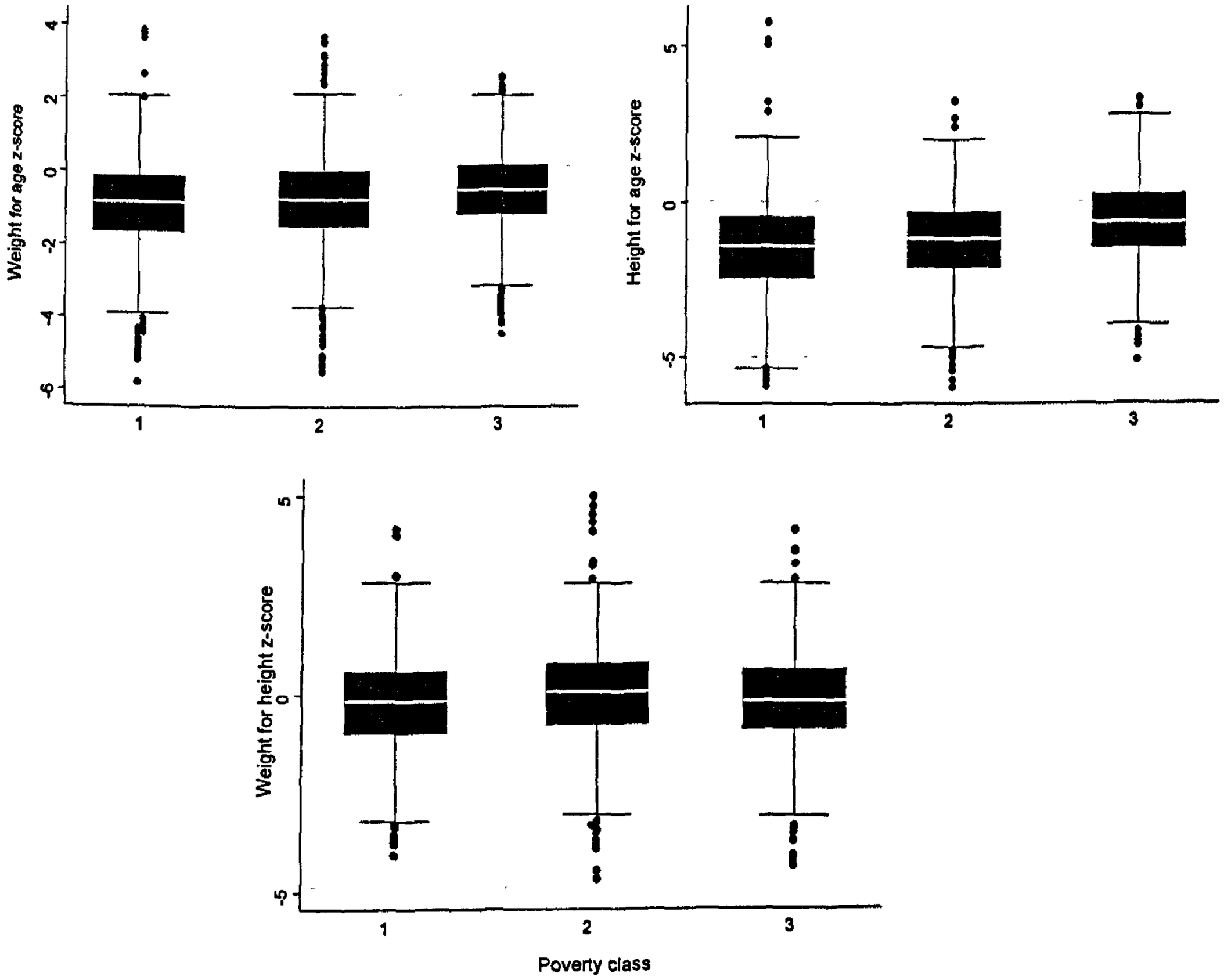


Figure 9.8 Association between malnutrition indicators and SES. X-axis denotes socio-economic classification - 1-poor, 2 – middle (top) and 3 - rich class (bottom).

As expected, crowding was associated with increased risk of disease both RSV specific and all-cause LRTI. It is hypothesized that this restriction of observed effect to disease is in line with the increase in inoculum in households of higher density known to produce greater disease [162, 294]. Having more children in the home is also expected to relate to an increased chance of virus introduction. The number of children in the home was highly correlated to number of persons in the home in this study population. Crowding also increases the potential for transmission taking place as a result of the increased chance and intensity of contact. Several recent studies (reviewed by Simoes [189]) observed a positive association between crowding and number of siblings and occurrence of RSV-LRTI. In this study, the age category (siblings < 6 years) of children in the home was found to be

associated with RSV disease (Table 9.2). This would imply that not only does the number of children in the home matter but also the age of these children. A possible explanation for this is that younger children are more susceptible to infection thus more likely to acquire the infection in the first place and in turn transmit to the young child.

We investigated several factors related to intensity of contact between other children in the home and the cohort child for all the three outcomes. An association was found between the intensity of contact (number of sibling sleeping in the same room) and increased risk of all-cause LRTI, Table 9.5. This correlates with the higher probability of transmission taking place due to prolonged exposure. A large number of studies as reviewed by Simoes [189] have shown an association between RSV-LRTI and number of people sleeping in the same room with the child or with siblings in school, none of these studies explored differences between sleeping in same room and more specifically in same bed as the child. Many studies [113, 162, 163, 170] have found an independently significant association between LRTI and increased numbers of others sharing the room. We investigated the potential effects of these different dynamics; number of siblings sleeping in the same house, room and bed. However, a similar association was not observed for RSV-LRTI possibly due to the reduced power to detect a difference given the small number of events.

The risks for RSV-LRTI can only really be interpreted in relation to those for RSVI in that it is not known if identified independent associations are for infection or disease, given that infection occurs in everyone. Thus factors identified to be independently associated with clinical RSVI were included in the model to determine factors that predict disease. It is clear that socio-economic classification (indicated by house and toilet type) plays an important role in the infection and disease in these data. In this study, living in a blocked

walled house with flush toilet ("richer" HH) was protective from RSVI whereas living in a block walled house with no toilet ("poor" HH) was independently associated with increased risk of RSV-LRTI. Though only of borderline significance, this finding is further supported by results indicating that children in households that receive financial assistance from an outside party/relative and those in households where the MIP is in trade had an increased risk of RSVI and disease respectively. This finding is inconsistent with findings from the study by Weber *et al* [194] where rather surprisingly case households more frequently had tap water and a flush toilet. They attribute this finding to differences in care seeking behavior by SES with parents of higher status being over represented in hosipitalised cases i.e. seeking treatment more effectively. In this study all children were under the same method of surveillance in addition to having their medical costs covered during the study. A rationalization of the effect of socioeconomics is that it is a determinant of factors that may influence health e.g. nutritional status and exposure to environmental smoke. Households that have a flush toilet and block wall in this population have a higher standard of living reflecting an economic advantage which is in turn reflected on their general health status.

Along the same line, having a hired care giver was independently associated with protection from RSVI. Very few households in this study population are able to afford hired help. Thus a child who is cared for by a house help is most likely living in a household that would be categorized as being better off financially. We assessed for an association between having a house help and a household's socio-economic classification and found that 80% of households that had hired help were classified as "rich" with the remaining 20% classified as "middle" class households. None of the poor households had hired help. Thus we can conclude that this observed protective effect is almost certainly a

proxy for socioeconomic effect. The pathway by which SES might affect risk of disease include: higher crowding index- more people fewer rooms –increases the likelihood of infection, poor general state of health a refection of poor quality of environment (poor sanitation, dirty water among other things [162]) and nutritional status (lower buying power-long term undernourishment). Also people of lower SES have been reported to have poor access to health care and also tend to seek healthcare late [295]. From this study it is not possible to discern what the SES effect is due to.

From this study it emerged that the age of the child's care taker plays a significant role in the risk of RSVI and RSV disease signified by an increased risk of infection and disease in children having older PCTs. There have been differing results on the effect of maternal age on LRTI, thus this finding though not surprising was unique for RSV infection and disease. In crude analysis of the Bostid studies [27] no consistent association was seen between maternal age and incidence of A(L)RI. With the exception of a few instances where the PCT was the grandmother, in most instances the PCT was the mother of the child. This finding may be related to or may be a proxy for family size with older mothers having had more children thus have generally more children to care for. It is also possible that older PCT are really grandparents looking after children whose mothers are working away, both of which are factors which may in turn affect the child's overall health translating into increased risk of RSV infection and disease.

Exposure to tobacco and to environmental smoke (cooking fuel) was associated with RSVI and all-cause LRTI respectively in these data, evidenced by increased risk of infection and disease. This result is similar to reports from other studies showing an excess risk of ARI and or RSV disease due to indoor air pollution and exposure to smoke[144, 164, 166, 169,

195]. It is widely known that environmental smoke contains significant amounts of pollutants (carbon monoxide, hydrocarbons) and also that the pathogenesis of ARI from biomass fuel has been shown to be similar to that of environmental tobacco smoke [296]. Still, the measurement of the independent variable, extent and length or intensity of exposure are issues to consider in assessing the effect. In this community female/maternal smoking is almost non-existent but exposure to cooking smoke is very prevalent. Most mothers tend to carry their young children on their backs while attending to household chores like cooking. Burning refuse was not associated with increased risk in any of the outcomes.

In this study, we observed evidence of an association between age and risk of LRTI. The risk of LRTI was decreased in the subgroup of children aged 6-11 months with an observed increase in children 12-17 months. A similar effect was not observed for RSVI and RSV-LRTI. Thus, this effect may be a reflection of disease by other causes (bacteria, other viruses) of LRTI in this population. This increase at 11 months is presumably co-incident with introduction of solid food and independent motility. Declining nutritional status (Chapter 8) is also an important influencing factor as the nutritional status of children in this population was observed to start deteriorating at around 9 months of age. Furthermore, except in the case of RSV (conflicting results), maternal antibody is usually known to be protective in the first few months of life, a possible explanation for the protective effect seen at 6-11 months waning as the child gets older. There was no evidence of association between age and RSV-LRTI. Several studies have reported a relationship between age and RSV disease with most disease occurring in infancy specifically in children <6 months [132, 133, 191-193]. It has been suggested [189, 294] that this is as a result of several effects; immature immune systems in younger child which improves with age, narrower

airways, hyper-responsiveness among others factors. The effect of age on disease may be limited to severe disease; published studies are cases of RSV-LRTI requiring hospitalization and are not similar to those in this study. We observed such an effect in this study looking at the incidence of severe RSV-LRTI by age. In this study, there appeared to be protection from severe RSV-LRTI with increasing age (Figure 9.1).

Educational level attained by mothers is also an important social factor determining child health. The mothers' educational level was associated with protection from all-cause LRTI. There are several ways through which maternal education is thought to influence child survival [264] and by extension the risk of disease: i) education alters the traditional balance of power within the family allowing the mother to take independent decisions concerning child health ii) Education also modifies beliefs and knowledge about disease causation, prevention and cure and this has an influence on health care practices iii) educated mothers bring their children much earlier and much more frequently for treatment iv) Schooling causes changes in behavior such as improved cleanliness which are likely to persist into adulthood with implications for the general health of the child [264]. A study done in Kenya to assess the knowledge, attitude and practices (KAP) of mothers regarding ARI found that formal education had a positive influence on the KAP of the mothers [297] In this study, there was a gradual increase in the level of effect with increasing number of years of education (Table 9.3), indicating that the better educated the mother is, the bigger the effect of education on reducing risk of disease, lending credence to some of the mechanisms of action of the education effect listed above. Nonetheless, it is worth mentioning that less than 5% of the mothers in this study had higher than a high school education with 25% having had no schooling at all.

Multiple births (triplets) were associated with increased risk of LRTI in this population. There were very few multiple births in these data, 2 triplets and 27 twin births. These children as is typical had lower birth weight, which is probably the underlying effect explained by this result. Several studies from different developing countries have shown low birth weight to be a risk factor for LRTI [166, 170, 179]. Nevertheless, the results reported are not consistent as some studies have refuted the effect of low birth weight [163, 170]. Low birth weight did not correlate with increased risk of LRTI in the multivariate model in these data, although there appeared to be a protective effect in heavier children in the univariate analyses for all the three outcomes. This effect however did not achieve significance. Another underlying factor is that children of multiple births are rarely carried to term and pre-maturity is a well known risk factors for ARI.

Having siblings in school doesn't seem to be a risk factor for infection or disease in these data despite studies [15, 16, 50, 144] showing that they are the most likely introducer of infection into the home. However, all these studies were done in developed countries which have markedly different household sizes and community structures, factors which may undoubtedly influence this effect considerably. An association between number of siblings in school and the risk of infection and disease was not observed in this study, although there appeared to be a borderline association between number of siblings <6yrs in school and increased risk of all-cause LRTI. The risk of transmission is thought to be higher in settings such as schools, therefore having older siblings who are more likely to be exposed to external infection and consequently facilitate infection in the birth cohort child, translating into higher risk of infection/disease in the cohort child. In fact in the household study, the rates of respiratory infection were found to be higher in pre-school children with siblings in school than in those without siblings in school (Chapter 5). However, the cohort

child in the household study was more often the primary case, possibly a result of undetected sub-clinical infection in other family members as has previously been discussed.

For all three outcome measures, the sex of the child had no observable effect despite the fact that younger male children are known to have more severe respiratory infection than girls due to their narrower and shorter airways [189, 298]. To investigate potential interaction between age and sex we looked at age-stratified incidence of infection by sex. There was no evidence for a difference in incidence of infection between males and females in the different age classes.

The differences observed in this and other studies may be related to the fact that: i) in this case one of the outcomes was clinical RSV infection. There have been few studies with RSVI as an outcome presumably because every child gets infected. Still, RSVI is the most important risk factors for RSV-LRTI. The comparison between predictors of clinical infection and disease in this study is new. It appears that the risk factors for each are markedly different with only two factors identified to be common to both: increasing age of PCT and house and toilet type (possible surrogate for SES). Malnutrition and crowding were associated RSV-LRTI. ii) Common knowledge on risk factors for RSV disease typically stems from studies done on hospitalized cases which are in all probability more serious requiring hospitalization, than most of the cases in these data. Published data from Kilifi showed that hospitalisations due to RSV-LRTI represent only a small fraction (perhaps 15%) of all severe RSV-LRTI in the community [5]; there were very many severe RSV cases that were not admitted adding weight to the theory that hospitalized cases maybe more severe. A comparison of results from the present study with those from other studies with differences in design may account for some of these differences. iii) there is

also a lack of precision in the specification of independent variables in the different studies with different definitions, sample collection methods and diagnostic methods used in different studies.

In this study nasal specimens were collected using the bulb washing method and tested by IFAT, the recommended method for diagnosis of RSV in samples from children. Nonetheless, diagnosis on the basis of immunofluorescent antibody test alone is not as sensitive as that by assay combinations [19] or by molecular methods. As previously discussed a study done in Kilifi showed the nasal wash bulb method to provide simple and effective alternatives to NPA (used in most hospital studies), with the NW being more acceptable and convenient for use in resource poor and home settings. In addition, there was no difference in the number of epithelial cells in the samples collected by the two methods [243]. The definition of LRTI used in this study was based on modified WHO guidelines specifically to capture LRTI in children <2 months and in cases of malaria co-infection. However, cases of RSV sero-conversion without an equivalent antigen positive result have been identified in this population (unpublished data-on going work) so it is possible and indeed probable that certain risk factors remain hidden due to reduced number of RSV-LRTI events.

In conclusion, the factors that increase the risk of RSV-LRTI and LRTI in the published literature generally vary possibly because of inherent differences in other factors like cultural and traditional practices, poverty levels, climatic conditions, political structure that influence health systems among other things. Thus there may not be a homogenous effect of risk factors on the incidence of respiratory tract infections the world over. In addition, some risk factors identified are not changeable: family size, exposure to siblings,

socioeconomic assets. Even though a causal relationship can only be demonstrated in a controlled trial, some of the risk factors identified can be modified by what would be considered relatively easy interventions. For instance, nutritional status of children can be improved by nutritional supplements (vitamin supplements or fortified milk) or encouraging longer breastfeeding ensured by better family planning methods i.e. well spaced out children which can be implemented through pre-natal education.

Chapter Ten

Overall Discussion

10.1 Introduction

The work presented in this thesis is an attempt to improve understanding of the infection dynamics of RSV within the community, and more specifically within the household. Most severe cases of RSV arise as primary infection in early infancy; for which the household contact is presumed to play a major role. Given the requirement for close contact to effect RSV transmission [299], the household within which intimate contact is undeniable presents a suitable unit of study to quantify community transmission rates. Presently, there is limited data on RSV in the resource poor settings and differences in family structure and contacts patterns may result in markedly different transmission dynamics to those identified in industrialized settings.

This thesis thus presents an account of highly detailed investigations of the transmission patterns of RSV within households in the setting of a rural developing country community: Kilifi District, coastal Kenya from 2003 to 2005. RSV infection was sought in both children and adults using clinical and serological surveillance. Routine active and passive visit records were analyzed to address each of the study objectives. Data on the follow-up of birth cohort children in the main project was used to investigate risk factors. A detailed discussion of objectives of the study, including comparison of findings with those published and the study limitations can be found at the end of each result chapter. The aim of this Chapter is to summarize findings and make recommendations for future work.

10.2 Findings and evidence to support findings

The source of infection to infants is not well documented. Thus, introduction and transmission of the virus within families was studied. This data is unique in resource poor settings and has not been collected anywhere since the 1970s. Surprisingly, the data indicate that introduction into the home was more often by the birth cohort child, although infants and young children mostly stay at home, so that *a priori* it would seem unlikely that they are the primary source of infection in households [15, 16, 61]. In fact, in households with documented secondary transmission in this study, older siblings were found to be the more likely introducers. It is epidemiologically plausible to expect that children in school are more likely to be infected externally (contact with other children) and in turn bring infection into the home. Moreover, infection rates were also found to be higher in pre-school children with siblings in school (discussed below). It is possible that an apparent introduction by a birth cohort child may in fact be a secondary case resulting from an inapparent asymptomatic primary case in a school-attending sibling. Due to the ethically constrained study design, where samples were collected only in the presence of symptoms, such a case would be unidentified.

Older children were also observed to transmit more effectively to younger household members relative to transmission among themselves; a likely reflection of the increased susceptibility of younger children. Overall, there was minimal secondary transmission observed to be taking place, but the number of people infected in a family in an epidemic was influenced by a household's socio-economic status (SES). It is not possible in the current study to say exactly how the SES effect actually manifests itself, but one could speculate that this effect could be due to higher crowding in poorer households. This would result in greater intensity of contact potentially increasing the likelihood of transmission

perhaps through multiple exposures or increased dose. However, family size was not shown to be related to size of outbreaks in infected households. Alternatively it may reflect the generally poorer health status of individuals in these families. These results are of course exclusive of possible missed asymptomatic infection, although appreciable levels of re-infection were documented.

It was hypothesized that re-infections have important influence on community transmission and persistence hence an attempt was made to estimate the age-stratified incidence of infection. This study demonstrates quite considerable infection levels in families, most of which is re-infection. Infections in older children and adults have largely been ignored by previous investigations. This study is unique in that previous studies have all been carried out in the developed world where social dynamics such as family sizes, living arrangements, and contact patterns are qualitatively different from those in developing countries. This confirms that infection in the older population can have important implications for transmission, infection, and disease in infants as possible reservoirs of infection.

Infection rates appear to be similar in the first 3 years of life and only start to decline after 3 years of age (Fig. 7.4). As expected, incidence of disease in this study was documented only in children younger than three years. Infection and disease rates were observed to be similar in both sexes. No cases of clinical infection were identified in adults (individuals over 15 years). In this study there were no cases identified of malaria (presence of parasites) co-infection in RSV cases. In areas of endemic malaria such as Kilifi, cases of malaria may be misdiagnosed as LRTI or ARI, hence giving an inaccurately high diagnosis

rate, so the fact that these patients were malaria negative gives strength to the diagnosis of respiratory disease.

There is a paucity of disease risk data for LDCs and the only other published study is by Nokes *et al* involving children in the first year of life from the same birth cohort as in the present study [5]. The household study was nested within the birth cohort project in which, for a subgroup of families, follow-up was extended to all family members (Chapter 3). As expected [12], risk of disease was found to be greater after primary infection than re-infection. Severe disease was observed to occur only during primary infection and only in children less than 18 months old. The risk of disease after re-infection was significantly higher in younger children. Interestingly, the highest risk of disease in this population was documented in the age group 12-17 months. This demonstrates that significant disease can occur following a primary infection even at this older age group, an observation which has not been previously documented.

Infection rates (ARI and RSV specific ARI) were highest in pre-school children with siblings in school. These children seem to be playing a part in bringing the infection into the home. This has important implications when considering possible vaccination strategies. Until recently [2], attempts at vaccinating infants have met several obstacles. The principal barriers have been the need to vaccinate in the first months of life, at a time of interference by maternal antibodies as well as immunologic immaturity observed in infants. Additionally, the formalin inactivated vaccine that was tested in the 1960's failed to protect and resulted in enhanced disease in vaccinated infants during naturally occurring RSV infections [91, 300]. Eliciting indirect disease protection in infants by vaccinating school children/siblings who are bringing infection into the home could potentially circumvent these problems. Reducing infection in school age children would reduce the

risk of infection in the vulnerable infants and young children and thus defer first infection to an older age where the risk of disease is significantly less – though as this study shows – not negligible. A comparable working model is the case with measles vaccine where the vaccine is given at 9 months. This does not directly protect the very young but might delay first infection by indirect protection conferred from less circulating virus in the community i.e. through the indirect effects of vaccination [86, 301]. Given that natural infection does not prevent re-infection, it is hypothesized that a vaccine would provide at least partial protection from re-infection but might need to be given repeatedly (e.g. every year) for a sustained effect sufficient to protect infants. Whether this would be acceptable (recipients of the vaccine would not be those to benefit most) is not entirely clear but has the potential of substantially reducing incidence of disease in infants, which may outweigh the cost of repeated vaccination. An additional problem with this approach is to provide protection in the very young (<3m), requiring attenuation to a degree that the vaccine immunogenity is compromised [2]; there is currently no attempt to produce a more immunogenic (and possibly reactogenic) vaccine for use in individuals who have already been exposed to a natural infection. A further problem would be that after a period of approximately 5 years of a programme vaccinating school-children to protect infants, if the programme is successful, individuals would potentially join the targeted age-groups without having been infected naturally.

Transmission is constrained by the duration of infectiousness *inter alia*. Thus an attempt was made to estimate the duration of viral shedding in this population. The longer the duration of shedding the greater the chance that transmission takes place (although account needs to be made of the shedding load which was not done in this study). Thus, any factor that affects this duration has important implications for disease incidence. The relationship

between past exposure, age, severity of infection, and sex and duration of shedding was explored. For some there was a suggestion of a risk (e.g. KM curves, Fig 6.6) but these were found to be statistically non-significant after accounting for some of the uncertainty in the data. The limitations of the data used (which may influence the results) have been discussed in Chapter 6. Further analysis of these data is required to tease out the possible relationships.

Finally, the main purpose of the risk factor survey was to determine the independent effects that certain factors have on the risk of RSV-LRTI in cohort children. To this effect, a total of 469 children followed up from birth were compared i.e. infected and non-infected for possible risk factors for RSV-LRTI. These were then compared with those factors identified to be associated with all-cause LRTI and thereby identify factors solely related to RSV disease. Only two factors were found to be independently associated with both RSV-LRTI and LRTI; stunting (a marker of long term under nutrition) and crowding (as indicated by number of children in the home). Those that were common to both RSV infection and RSV-LRTI and associated with increased risk of infection and disease were increased age of the PCT and socio-economic status. All other independent factors were unique to each outcome. House type and sanitation was associated with increased risk of RSV disease but not with LRTI as was having a sibling under the age of 6 years. The only risk factor that was unique to infection with RSV was the number of smokers in the home. Age at infection and multiple births were found to be associated with increased risk of LRTI while a higher level of maternal education was associated with protection.

Crowding (discussed above) and nutritional status are both directly influenced by SES; nutritional status is affected by a families' long term ability to purchase food. Educational

level can also be linked to SES. It is speculated that education may permit better employment opportunity facilitating stability through a regular income. A study done in Kilifi [295] observed that poor families tended to access health services later, i.e. they had other pressing priorities or little or no disposable income and faired worse than their wealthier counterpart in the event of an illness episode in the family. Nonetheless, to demonstrate a causal relationship between an observed, putative risk factor and disease outcome, an intervention through a controlled trial or by comparison would be necessary. However, nutritional supplementation and education are both already acknowledged to have significant positive benefits on all outcomes (not just RSV associated disease), so that such an intervention study is unlikely to be possible just based on RSV.

In this study there were three epidemics observed in the household cohort, one at recruitment and two occurring in close succession almost a year later (Fig. 2.3), seemingly having a similar pattern to that seen in Finland [44]. However, more surveillance data is required to state this conclusively. The peak season for RSV was correlated with the absence of rainfall (Fig. 3.10). No re-infections were identified during the inter-epidemic period leaving the question of persistence unresolved. Although this data adds to the global knowledge on seasonality of RSV, it does not provide an explanation for the observed seasonal dynamics.

10.3 Generalisability of study findings

It is 50 years since RSV was discovered and as yet, information on RSV from developing countries is remarkably lacking. The WHO recognised the importance of RSV disease burden data in developing countries, supporting a generic surveillance protocol and funding a group of studies on RSV disease burden [11, 117]. The present study provides data that

will contribute further to the knowledge base on RSV in developing countries and worldwide. Such a study is desirable as it provides information about the spread of RSV in the community, information that is vital in relation to the future of vaccine trials and interventions [2]. These series of studies have provided reliable estimates (albeit underestimated) of the amount of infection in certain age groups in the community and has shown that RSV poses a significant burden in this rural African population – not just to the infant but older age groups. This is the first such study from a developing country.

Risk factor data on RSV from developing countries is also limited [126, 194]. These results thus add to the body of knowledge on risk factors for RSV disease. Different study designs and population dynamics will ultimately affect observed findings. Nevertheless, other studies in similar settings such as this (rural community in a developing country) should ideally provide comparable results.

A series of studies conducted in the 1970s and early 1980s in USA contributed to the understanding of the pattern of RSV shedding in hospitalized young infants without underlying conditions [61-63] and in the community [15, 59]. It is only in these studies that the duration and titres shed by infants and young children have been classically reported. Thus this study of non-hospitalised young and older children in relation to history of infection provides important additional information relative to the data available.

10.4 Study Limitations

Generally, serological information on infections, based on repeated serum or oral-fluid sampling, would have greatly improved the precision of the incidence estimates. In particular, supporting data from serological confirmed cases would have overcome the difficulty of collecting nasal washes and assay sensitivity in older children and adults. Collection of oral fluid is arguably a more acceptable alternative to serum. In particular, it is felt that results from the oral-fluid assay would have enabled a truer estimation of incidence and made possible the documentation of transmission in the non-epidemic period possibly shedding light on viral persistence. In addition, it would probably have influenced the apparent result that primary cases on infants fed infection in the household rather than re-infections from older siblings. However, a qualitative alteration of our findings of higher incidence in 12-18m olds and risk of disease estimates is unlikely as most primary infections are symptomatic.

There were additional limitations of the surveillance methods in general. A severe limitation was that samples were only collected from individuals with symptoms. Without collecting specimens frequently and routinely in the absence of symptoms, there was a high potential for missing many infections, and the less frequent sampling of oral fluid (or serum) for serological confirmation has the drawback of not pin-pointing when the infection arose precisely in time and so cannot be used to identify transmission routes, thus the potential beneficial ability of active and passive methods to identify all cases at the time they arise. However, this is not necessary to determine infection in the inter-epidemic period but is crucial for transmission network studies. Another limitation of the present study was our inability to collect nasal specimens from adults pointing to the need for a more suitable method of collecting nasal specimens. Following on from this, a less

'invasive' yet sensitive method that would allow collection of samples irrespective of symptoms would be extremely beneficial to future studies.

Prolonged shedders (shedding virus for more than 2 weeks) have been identified in other studies [302], but are likely to have been missed in this study because the duration of shedding was curtailed by taking the first negative test after positive as cessation of shedding interval. For those individuals from whom samples were collected until two consecutive negatives, only 4.9 % were identified as being falsely negative, so this is unlikely to seriously influence our estimates of duration of infection, but would potentially result in missing a small number of individuals with prolonged shedding, which might be discontinuous or variable. RSV infections have been detected throughout the year in HIV infected children [41] suggesting the possibility of prolonged shedding in immuno-compromised individuals [302]. The HIV status of children included in this study was not available, in part because HIV VCT (voluntary counseling and testing), hospital Diagnostic CT, and accompanying support infrastructure were at an early stage of implementation. Moreover, possibly more objective data by which to judge severity may be necessary although this could prove difficult in resource poor settlings.

10.5 Recommendations for Further work

There are several areas relating to re-infection and community transmission which still need to be explored. First it may prove necessary in future studies to visit households at shorter intervals possibly twice a week i.e. every 3-4 days. Furthermore, the question still remains open regarding how much infection is actually taking place in adults in this community and in turn relating this to infection in children. Given the improvement observed in results

from the present serum ELISA the possibility of reviewing the oral-fluid assay with better results seems likely and can be used to shed light on this issue.

Genotypic profiling of the genotypes circulating within the family members during epidemics would provide information on who infects whom within the household structure. In no instance has the transmission within families been investigated with respect to the different genotypes circulating. It would be important therefore to establish whether the strains circulating in other members of the family especially the school aged siblings are similar to that found in the infant thus confirming possible association for the source of infection. Thus future work would need to include this.

Future studies to determine the duration of shedding would need to continue sampling for at least one more occasion after the first negative sample. Additionally, there might be an inherent bias in the way the data was collected against detecting shorter shedding periods. Thus the possibility of regular sampling with shorter interval even in the absence of respiratory symptoms should be explored. Regular sampling in the absence of symptoms would also identify infections that may not elicit symptoms. Although this may be difficult to justify it seems necessary. Lastly, it is possible that transmission and sequelae of infection depend in part on the amount of RSV inoculum transmitted [303]. Thus, future reports on duration of RSV shedding should include a quantitation of the amount of virus shed.

Lastly, there has been so little success in the current main approach to vaccination i.e. direct protection of infants in first few months. Other strategies such as maternal antibody boosting to provide longer duration of protection in first 6 months need exploration.

However, given the rate of catabolism of maternal antibodies, and resultant half life of 1 to 3 months [88, 89, 304, 305], this strategy may not substantially decrease the proportion at risk in first 6 months. Alternative strategies may therefore include immunization of those with partial immunity e.g. school children or all children within households of a newborn (discussed above) delaying the first infection thus potentially reducing the risk of severe disease in early life.

References:

[1] Belshe RB, Newman FK, Anderson EL, Wright PF, Karron RA, Tollefson S, et al. Evaluation of combined live, attenuated respiratory syncytial virus and parainfluenza 3 virus vaccines in infants and young children. J Infect Dis. 2004 Dec 15;190(12):2096-103.

[2] Karron RA, Wright PF, Belshe RB, Thumar B, Casey R, Newman F, et al. Identification of a recombinant live attenuated respiratory syncytial virus vaccine candidate that is highly attenuated in infants. J Infect Dis. 2005 Apr 1;191(7):1093-104.

[3] Wright PF, Karron RA, Belshe RB, Thompson J, Crowe JE, Jr., Boyce TG, et al. Evaluation of a live, cold-passaged, temperature-sensitive, respiratory syncytial virus vaccine candidate in infancy. J Infect Dis. 2000 Nov;182(5):1331-42.

[4] Sutmoller F, Ferro ZP, Asensi MD, Ferreira V, Mazzei IS, Cunha BL. Etiology of acute respiratory tract infections among children in a combined community and hospital study in Rio de Janeiro. Clin Infect Dis. 1995 Apr;20(4):854-60.

[5] Nokes DJ, Okiro EA, Ngama M, White LJ, Ochola R, Scott PD, et al. Respiratory syncytial virus epidemiology in a birth cohort from Kilifi district, Kenya: infection during the first year of life. J Infect Dis. 2004 Nov 15;190(10):1828-32.

[6] Borrero I, Fajardo L, Bedoya A, Zea A, Carmona F, de Borrero M, F. Acute respiratory tract infections among a birth cohort of children from Cali, Colombia, who were studied through 17 months of age. . Reviews of Infectious Diseases. 1990;12:Suppl 8:S950-6.

[7] Berman S, Duenas A, Bedoya A, Constain V, Leon S, Borrero I, et al. Acute lower respiratory tract illnesses in Cali, Columbia: a two year ambulatory study. Pediatrics. 1983;71:210-8.

[8] Djelantik IG, Gessner BD, Soewignjo S, Steinhoff M, Sutanto A, Widjaya A, et al. Incidence and clinical features of hospitalization because of respiratory syncytial virus lower respiratory illness among children less than two years of age in a rural Asian setting. Pediatr Infect Dis J. 2003;22(2):150-7.

[9] Weber MW, Milligan P, Sanneh M, Awemoyi A, Dakour R, Schneider G, et al. An epidemiological study of RSV infection in the Gambia. Bull World Health Organ. 2002;80(7):562-8.

[10] Dagan R, Landau D, Haikin H, Tal A. Hospitalization of Jewish and Bedouin infants in southern Israel for bronchiolitis caused by respiratory syncytial virus. Pediatr Infect Dis J. 1993 May;12(5):381-6.

[11] Robertson SE, Roca A, Alonso P, Simoes EA, Kartasasmita CB, Olaleye DO, et al. Respiratory syncytial virus infection: denominator-based studies in Indonesia, Mozambique, Nigeria and South Africa. Bull World Health Organ. 2004 Dec;82(12):914-22.

[12] Glezen W, Taber L, Frank A, Kasel J. Risk of primary infection and reinfection with respiratory syncytial virus. American Journal of Diseases of Children. 1986;140:543-6.

[13] Fisher RG, Gruber WC, Edwards KM, Reed GW, Tollefson SJ, Thompson JM, et al. Twenty years of outpatient respiratory syncytial virus infection: a framework for vaccine efficacy trials. Pediatrics. 1997;99(2):E7.

[14] Hall C, Douglas RJ, Geiman J. Possible transmission by fomites of respiratory syncytial virus. Journal of infectious Diseases. 1980;141:98-102.

[15] Hall CB, Geiman JM, Biggar R, Kotok DI, Hogan PM, Douglas GR, Jr. Respiratory syncytial virus infections within families. N Engl J Med. 1976;294(8):414-9.

[16] Monto A, Sook K. The Tecumesh study of respiratory ilness. III. incidence and periodicity of respiratory syncytial virus and Mycoplasma pneumoniae infections. American Journal of Epidemiology. 1971;94:290-301.

[17] Hall CE, Brandt CD, Frothingham TE, Spigland I, Cooney MK, Fox JP. The virus watch program: a continuing surveillance of viral infections in metropolitan New York families IX. A comparison of infections with several respiratory pathogens in New York and New Orleans families. Am J Epidemiology. 1971;94:436-54.

[18] Beem M. Repeated infections with respiratory syncytial virus. Journal of Immunology. 1967;98:1115-22.

[19] Henderson F, Collier A, Clyde WJ, Denny F. Respiratory-syncytial-virus infections, reinfections and immunity. A prospective, longitudinal study in young children. New England Journal of Medicine. 1979;300:530-4.

[20] Wilson S, Roberts K, Hammond K, Ayres J, Cane P. Estimation of incidence of respiratory syncytial virus infections in school children using salivary antibodies. Journal of Medical Virology. 2000;61:81-4.

[21] Mulholland K. Global Burden of Acute Respiratory Infections in Children: implications for interventions
Pediatric Pulmonology. 2003;36(6):469-74.

[22] Campbell H. Acute respiratory infection: a global challenge. Arch Dis Child. 1995 Oct;73(4):281-3.

[23] Garenne M RC, Campbell H. The magnitude of mortality from acute respiratory infections in children under 5 years in developing countries. World Health Stat Q. 1992 1992; 45((2-3)):180-91.

[24] Williams BG, Gouws E, Boschi-Pinto C, Bryce J, Dye C. Estimates of world-wide distribution of child deaths from acute respiratory infections. Lancet Infect Dis. 2002 Jan;2(1):25-32.

[25] Bryce J B-PC, Shibuya K, Black RE, Group. WCHER. WHO estimates of the causes of death in children. Lancet. 2005 2005 Mar 26-Apr 1;365(9465):1147-52.

[26] Kenya Ministry of Health. Health management information systems: Report for the 1996 to 1999 period. In: Health KMo, ed.: Kenya Ministry of Health 2001.

[27] Selwyn B. The epidemiology of acute respiratory tract infection in young children: comparison of findings from several developing countries. Reviews of Infectious Diseases. 1990;12 (supplement 8):S870-S88.

[28] Weber M, Mulholland E, Greenwood B. Respiratory syncytial virus infection in tropical and developing countries. Tropical Medicine and Internationl Health. 1998;3(4):268-80.

[29] Moura FE, Borges LC, Portes SA, Ramos EA, Siqueira MM. Respiratory syncytial virus infections during an epidemic period in Salvador, Brazil. Viral antigenic group analysis and description of clinical and epidemiological aspects. Mem Inst Oswaldo Cruz. 2003 Sep;98(6):739-43.

[30] Chanock RM, Parrott RH, Vargosko AJ, Kapikian AZ, Knight V, Johnson KM. Acute respiratory diseases of viral etiology. IV. Respiratory syncytial virus. Am J Public Health. 1962 Jun;52:918-25.

[31] Beem M, Wright FH, Hamre D, Egerer R, Oehme M. Association of the chimpanzee coryza agent with acute respiratory disease in children. N Engl J Med. 1960 Sep 15;263:523-30.

[32] McClelland L, Hilleman MR, Hamparian VV, Ketler A, Reilly CM, Cornfeld D, et al. Studies of acute respiratory illnesses caused by respiratory syncytial virus. 2. Epidemiology and assessment of importance. N Engl J Med. 1961 Jun 8;264:1169-75.

[33] Chanock RM, Kim HW, Vargosko AJ, Deleva A, Johnson KM, Cumming C, et al. Respiratory syncytial virus. I. Virus recovery and other observations during 1960 outbreak

of bronchiolitis, pneumonia, and minor respiratory diseases in children. Jama. 1961 May 27;176:647-53.

[34] Lina B, Valette M, Foray S, Luciani J, Stagnara J, See DM, et al. Surveillance of community-acquired viral infections due to respiratory viruses in Rhone-Alpes (France) during winter 1994 to 1995. J Clin Microbiol. 1996 Dec;34(12):3007-11.

[35] Kim H, Arrobio J, Brandt C, Jefries B, Pyles G, Reid J, et al. Epidemiology of respiratory syncytial virus infection in Washington, D.C. I Importance of the virus in different respiratory tract disease syndromes and temporal distribution of infection. American Journal of Epidemiology. 1973;98:216-25.

[36] John TJ, Cherian T, Steinhoff MC, Simoes EA, John M. Etiology of acute respiratory infections in children in tropical southern India. Rev Infect Dis. 1991 May-Jun;13 Suppl 6:S463-9.

[37] Wright AL, Taussig LM, Ray CG, Harrison HR, Holberg CJ. The Tucson Children's Respiratory Study. II. Lower respiratory tract illness in the first year of life. Am J Epidemiol. 1989 Jun;129(6):1232-46.

[38] Nokes D. Respiratory syncytial virus disease burden in the developing world. . In: Cane PA, ed. *In Perspectives in Medical Virology: Respiratory Syncytial Virus* Elsevier 2006:183-230.

[39] Stensballe LG, Devasundaram JK, Simoes EA. Respiratory syncytial virus epidemics: the ups and downs of a seasonal virus. Pediatr Infect Dis J. 2003 Feb;22(2 Suppl):S21-32.

[40] Karron RA, Singleton RJ, Bulkow L, Parkinson A, Kruse D, DeSmet I, et al. Severe respiratory syncytial virus disease in Alaska native children. RSV Alaska Study Group. J Infect Dis. 1999;180(1):41-9.

[41] Madhi SA, Schoub B, Simmank K, Blackburn N, Klugman KP. Increased burden of respiratory viral associated severe lower respiratory tract infections in children infected with human immunodeficiency virus type-1. J Pediatr. 2000 Jul;137(1):78-84.

[42] Cane P, Matthews D, Pringle C. Analysis of respiratory syncytial virus strain variation in successive epidemics in one city. J Clin Microbiol. 1994;32:1-4.

[43] Hall C, Walsh E, Schnabel K, CE L, McConnochie K, Hildreth S, et al. Occurrence of groups A and B respiratory syncytial virus over 15 years: associated epidemiologic and clinical characteristics in hospitalized and ambulatory children. Journal of Infectious Diseases. 1990;162:1283-90.

[44] Waris M. Pattern of respiratory syncytial virus epidemics in Finland: two-year cycles with alternating prevalence of groups A and B. Journal of Infectious Diseases. 1991;163:464-9.

[45] Cherian T, Simoes EA, Steinhoff MC, Chitra K, John M, Raghupathy P, et al. Bronchiolitis in tropical south India. Am J Dis Child. 1990 Sep;144(9):1026-30.

[46] Hijazi Z, Pacsa A, Eisa S, el Shazli A, Abd el-Salam R, el-Gharbawy F. Respiratory syncytial virus infections in children in a desert country. Pediatr Infect Dis J. 1995 Apr;14(4):322-4.

[47] Weissenbacher M, Carballal G, Avila M, Salomon H, Harisiadi J, Catalano M, et al. Etiologic and clinical evaluation of acute lower respiratory tract infections in young Argentinian children: an overview. Rev Infect Dis. 1990 Nov-Dec;12 Suppl 8:S889-98.

[48] McAnerney JM, Johnson S, Schoub BD. Surveillance of respiratory viruses. A 10-year laboratory-based study. S Afr Med J. 1994 Aug;84(8 Pt 1):473-7.

[49] Waris M, White LJ. Seasonality of respiratory syncytial virus infection. Clin Infect Dis. 2006 Aug 15;43(4):541.

[50] Berglund B. Respiratory syncytial virus infections in families. A study of family members of children hospitalized for acute respiratory disease. Acta Paediatr Scand. 1967 Jul;56(4):395-404.

[51] Collins PL MK, Chanock RM. . Respiratory Syncytial Virus. In: Fields BN KD, Howley PM, ed. *Fields Virology 4th ed.* Philadelphia: Lippincott-Raven 1995:1443-85.

[52] Collins CL, Pollard AJ. Respiratory syncytial virus infections in children and adults. J Infect. 2002 Jul;45(1):10-7.

[53] Falsey AR, Walsh EE. Respiratory syncytial virus infection in adults. Clin Microbiol Rev. 2000 Jul;13(3):371-84.

[54] http://www.cdc.gov/ncidod/dvrd/revb/respiratory/rsvfeat.htm. Respiratory Syncytial Virus. 2005 [cited 2005 Feb 2005]; Available from: http://www.cdc.gov/ncidod/dvrd/revb/respiratory/rsvfeat.htm

[55] Collins PL, R. M. Chanock, et al. . Respiratory syncytial virus. In: Howley. DMKaPM, ed. *Fields Virology*. Philadelphia: Lippincott-Raven 2001:1443-85.

[56] Hall CB. The nosocomial spread of respiratory syncytial viral infections. Annu Rev Med. 1983;34:311-9.

[57] Johnson KM, Chanock RM, Rifkind D, Kravetz HM, Knight V. Respiratory syncytial virus. IV. Correlation of virus shedding, serologic response, and illness in adult volunteers. Jama. 1961 May 27;176:663-7.

[58] Kravetz HM, Knight V, Chanock RM, Morris JA, Johnson KM, Rifkind D, et al. Respiratory syncytial virus. III. Production of illness and clinical observations in adult volunteers. Jama. 1961 May 27;176:657-63.

[59] Frank AL, Taber LH, Wells CR, Wells JM, Glezen WP, Paredes A. Patterns of shedding of myxoviruses and paramyxoviruses in children. J Infect Dis. 1981 Nov;144(5):433-41.

[60] White LJ, Waris M, Cane PA, Nokes DJ, Medley GF. The transmission dynamics of groups A and B human respiratory syncytial virus (hRSV) in England & Wales and Finland: seasonality and cross-protection. Epidemiol Infect. 2005 Apr;133(2):279-89.

[61] Hall CB. The shedding and spreading of respiratory syncytial virus. Pediatr Res. 1977 Mar;11(3 Pt 2):236-9.

[62] Hall CB, Douglas RG, Jr., Geiman JM. Respiratory syncytial virus infections in infants: quantitation and duration of shedding. J Pediatr. 1976;89(1):11-5.

[63] Hall C, Douglas, RG Jr and Geiman, JM. Quantitative shedding patterns of RSV in infants. J Infect Dis. 1975 Aug;132(2):151-6.

[64] Waris M, O M, Mufson M, Ruuskanen O, Halonen P. Shedding of Infectious Virus and Virus Antigen During Acute Infection with Respiratory Syncytial Virus. Journal of Medical Virology. 1992;38:111-6.

[65] Hall CB, Long CE, Schnabel KC. Respiratory syncytial virus infections in previously healthy working adults. Clin Infect Dis. 2001;33(6):792-6.

[66] Peret T, Hall C, Schnabel K, Golub J, Anderson L. Circulation patterns of genetically distinct group A and B strains of human respiratory syncytial virus in a community. Journal of General Virology. 1998;79:2221-9.

[67] Coggins WB, E. J. Lefkowitz, and W. M. Sullender. Genetic variability among group A and group B respiratory syncytial viruses in a children's hospital. J Clin Microbiol. 1998;36:3552–7.

[68] Hacking DaJH. Respiratory Syncytial Virus - Viral Biology and the Host Response. Journal of Infection. 2002;45:18-24.

[69] Olmstead R. A. ea. ? Proc Natl Acad Sci USA. 1986;83:7462-6.

[70] Cherrie AH, K. Anderson, G. W. Wertz, and P. J. M. Openshaw. Human cytotoxic T cells stimulated by antigen on dendritic cells recognize the N, SH, F, M, 22K, and 1b proteins of respiratory syncytial virus. J Virol 1992;66:2102-10.

[71] Karron RA, Buonagurio DA, Georgiu AF, Whitehead SS, Adamus JE, Clements-Mann ML, et al. Respiratory syncytial virus (RSV) SH and G proteins are not essential for viral replication in vitro: clinical evaluation and molecular characterization of a cold-passaged, attenuated RSV subgroup B mutant. Proc Natl Acad Sci U S A. 1997 Dec 9;94(25):13961-6.

[72] Sullender WM, L. J. Anderson, M. A. Mufson, and G. W. Wertz. Genetic diversity of the attachment protein of subgroup B respiratory syncytial viruses. J Virol. 1991;65:5425–34.

[73] Sanz MC, O. M. Kew, and L. J. Anderson. Genetic heterogeneity of the attachment glycoprotein G among group A respiratory syncytial viruses. Virus Res. 1994;33:203–17.

[74] Sullender WM. Respiratory Syncytial Virus Genetic and Antigenic Diversity. Clinc Microbiol Rev. 2000;13(1):1-15.

[75] Cane PA. Molecular epidemiology of respiratory syncytial virus. Rev Med Virol. 2001;11(2):103-16.

[76] Sullender WM, Mufson MA, Prince GA, Anderson LJ, Wertz GW. Antigenic and genetic diversity among the attachment proteins of group A respiratory syncytial viruses that have caused repeat infections in children. J Infect Dis. 1998 Oct;178(4):925-32.

[77] Handforth J, Friedland JS, Sharland M. Basic epidemiology and immunopathology of RSV in children. Paediatr Respir Rev. 2000 Sep;1(3):210-4.

[78] Anderson LJ, Hierholzer JC, Tsou C, Hendry RM, Fernie BF, Stone Y, et al. Antigenic characterization of respiratory syncytial virus strains with monoclonal antibodies. J Infect Dis. 1985;151(4):626-33.

[79] Hendry RM, A. L. Talis, E. Godfrey, L. J. Anderson, B. F. Fernie, and K. McIntosh. Concurrent circulation of antigenically distinct strains of respiratory syncytial virus during community outbreaks. J Infect Dis. 1986;153:291–7.

[80] Mufson M, Orvell C, Rafnar B, Norrby E. Two distinct subtypes of human respiratory syncytial virus. Journal of General Virology. 1985;66:2111-24.

[81] Mufson MA, R. B. Belshe, C. Orvell, and E. Norrby. Respiratory syncytial virus epidemics: variable dominance of subgroups A and B strains among children, 1981–1986. J Infect Dis. 1988;157:143–8.

[82] Venter M, Madhi SA, Tiemessen CT, Schoub BD. Genetic diversity and molecular epidemiology of respiratory syncytial virus over four consecutive seasons in South Africa: identification of new subgroup A and B genotypes. J Gen Virol. 2001 Sep;82(Pt 9):2117-24.

[83] Scott PD, Ochola R, Ngama M, Okiro EA, Nokes DJ, Medley GF, et al. Molecular epidemiology of respiratory syncytial virus in Kilifi district, Kenya. J Med Virol. 2004 Oct;74(2):344-54.

[84] Reyes M, Eriksson M, Bennet R, Hedlund KO, Ehrnst A. Regular pattern of respiratory syncytial virus and rotavirus infections and relation to weather in Stockholm, 1984--1993. Clin Microbiol Infect. 1997 Feb;3(6):640-6.

[85] Weber A, Weber M, Milligan P. Modeling epidemics caused by respiratory syncytial virus (RSV). Math Biosci. 2001 Aug;172(2):95-113.

[86] Anderson RM, May R. Infectious Diseases of Humans: Dynamics and Control Oxford Science Publications 1992.

[87] Fine PE, Clarkson JA. Measles in England and Wales--I: An analysis of factors underlying seasonal patterns. Int J Epidemiol. 1982 Mar;11(1):5-14.

[88] Brandenburg AH, Groen J, van Steensel-Moll HA, Claas EC, Rothbarth PH, Neijens HJ, et al. Respiratory syncytial virus specific serum antibodies in infants under six months of age: limited serological response upon infection. J Med Virol. 1997 May;52(1):97-104.

[89] Cox M, Azevedo, RS., Cane, PA., Massad, E., Medley, GF. Seroepidemiological Study of Respiratory Syncytial Virus in Sao Paulo State, Brazil. Journal of Medical Virology. 1998 Apr 6;55(3):1.

[90] Parrott R, Kim H, Arrobio J, Hodes D, Murphy B, Brandt C, et al. Epidemiology of respiratory syncytial vrius in fection in Washington, D.C. II Infection and disease with respect to age, immunologic status, race and sex. American Journal of Epidemiology. 1973;98:289-300.

[91] Chin J, Magoffin RL, Shearer LA, Schieble JH, Lennette EH. Field evaluation of a respiratory syncytial virus vaccine and a trivalent parainfluenza virus vaccine in a pediatric population. Am J Epidemiol. 1969 Apr;89(4):449-63.

[92] Kasel JA, Walsh EE, Frank AL, Baxter BD, Taber LH, Glezen WP. Relation of serum antibody to glycoproteins of respiratory syncytial virus with immunity to infection in children. Viral Immunol. 1987;1(3):199-205.

[93] Ogilvie MM, Vathenen AS, Radford M, Codd J, Key S. Maternal antibody and respiratory syncytial virus infection in infancy. J Med Virol. 1981;7(4):263-71.

[94] Glezen W, Paredes A, Allison J, Taber L, Frank A. Risk of respiratory syncytial virus infection for infants from low-income families in relationhsip to age, sex, ethnic group, and maternal antibody level. Journal of Pediatrics. 1981;98:708-15.

[95] Bruhn FaAY. Respiratory syncytial virus in early infancy: circulating antibody and the severity of infection. Am J Dis Child. 1977;131:145-8.

[96] McCarthy C, Hall C. Respiratory Syncytial Virus: Concerns and Control. Pediatrics in Review. 2003 September;24(9):301-8.

[97] Groothuis JR, Simoes EA, Levin MJ, Hall CB, Long CE, Rodriguez WJ, et al. Prophylactic administration of respiratory syncytial virus immune globulin to high-risk infants and young children. The Respiratory Syncytial Virus Immune Globulin Study Group. N Engl J Med. 1993 Nov 18;329(21):1524-30.

[98] Walsh EE, Falsey AR. Humoral and mucosal immunity in protection from natural respiratory syncytial virus infection in adults. J Infect Dis. 2004 Jul 15;190(2):373-8.

[99] Mills Jt, Van Kirk JE, Wright PF, Chanock RM. Experimental respiratory syncytial virus infection of adults. Possible mechanisms of resistance to infection and illness. J Immunol. 1971;107(1):123-30.

[100] Watt PJ, Robinson BS, Pringle CR, Tyrrell DA. Determinants of susceptibility to challenge and the antibody response of adult volunteers given experimental respiratory syncytial virus vaccines. Vaccine. 1990 Jun;8(3):231-6.

[101] Hall C, Walsh E, Long C, Schnabel K. Immunity to and frequency of reinfection with respiratory syncytial virus. Journal of Infectious Diseases. 1991;163:693-8.

[102] Simoes EA. Respiratory syncytial virus infection. Lancet. 1999;354(9181):847-52.

[103] Beem M, Egerer, R. and Anderson, J. Respiratory syncytial virus neutralizing antibodies in persons residing in Chicago, Illinois. Pediatrics. 1964;34:761-70.

[104] Piedra PA, Jewell AM, Cron SG, Atmar RL, Glezen WP. Correlates of immunity to respiratory syncytial virus (RSV) associated-hospitalization: establishment of minimum protective threshold levels of serum neutralizing antibodies. Vaccine. 2003 Jul 28;21(24):3479-82.

[105] Cane P, Thomas H, Simpson A, Evans J, Hart C, Pringle C. Analysis of the human serological immune response to a variable region of the attachment (G) protein of the

respiratory syncytial virus during primary infection. Journal of Medical Virology. 1996;48:252-61.

[106] Cane P. Analysis of linear epitopes recognized by the primary human antibody response to variable region of the attachment (G) protein of respiratory syncytial virus. J Med Virol 1997. 1997;51:297±304.

[107] Anderson LJ, Hendry R. M., Pierik L.T., et al. Multicenter Study of strains of respiratory syncytial virus. J Infect Dis. 1991;163:687-92.

[108] Storch GA, C. B. Hall, L. J. Anderson, C. S. Park, and D. E. Dohner. Antigenic and nucleic acid analysis of nosocomial isolates of respiratory syncytial virus. J Infect Dis. 1993;167:562-6.

[109] Mufson M, Belshe R, Orvell C, Norrby E. Subgroup characteristics of respiratory syncytial virus strains recovered from children with two consecutive infections. J Clin Microbiol. 1987;25:1535-9.

[110] Scott PD, Ochola R, Ngama M, Okiro EA, James Nokes D, Medley GF, et al. Molecular analysis of respiratory syncytial virus reinfections in infants from coastal Kenya. J Infect Dis. 2006 Jan 1;193(1):59-67.

[111] Ramillo. An update on respiratory syncytial virus. Pediatric Infectious Disease Journal. 2002;21:453.

[112] Avendano LF, Larranaga C., Palomino M.A. Community- and hospital-acquired respiratory syncytial virus infections in Chile. Pediatr Infect Dis J. 1991;10:564-8.

[113] Holberg CJ, Wright A.L., Matinez F.D. Risk factors for respiratory syncytial virus-associated lower respiratory illness in the first year of life. Am JEpi. 1991;133:1135-51.

[114] Ukkonen P, Hovi T, von Bonsdorff CH, Saikku P, Penttinen K. Age-specific prevalence of complement-fixing antibodies to sixteen viral antigens: a computer analysis of 58,500 patients covering a period of eight years. J Med Virol. 1984;13(2):131-48.

[115] Monto AS, Ullman BM. Acute respiratory illness in an American community. The Tecumseh study. Jama. 1974;227(2):164-9.

[116] Weber MW DR, Usen S, Schneider G, Adegbola RA, Cane P, Jaffar S, Milligan P, Greenwood BM, Whittle H, Mulholland EK. The clinical spectrum of respiratory syncytial virus disease in The Gambia. Pediatr Infect Dis J. 1998 Mar;17(3):224-30.

[117] Wright P, Cutts F. Generic protocol to examine the incidence of lower respiratory infection (LRI) due respiratory syncytial virus (RSV) in children less than five years of age. Geneva: World Health Organization.; 2000.

[118] Hall C. Respiratory syncytial virus. 3 ed. Philadelphia: WB Saunders Company 1992.

[119] Beem MO. Acute respiratory illness in nursery school children: a longitudinal study of the occurrence of illness and respiratory viruses. Am J Epidemiol. 1969;90(1):30-44.

[120] Johnson KM, Bloom HH, Mufson MA, Chanock RM. Natural reinfection of adults by respiratory syncytial virus. Possible relation to mild upper respiratory disease. N Engl J Med. 1962 Jul 12;267:68-72.

[121] Dowell SF, Anderson LJ, Gary HE, Jr., Erdman DD, Plouffe JF, File TM, Jr., et al. Respiratory syncytial virus is an important cause of community-acquired lower respiratory infection among hospitalized adults. J Infect Dis. 1996;174(3):456-62.

[122] Hall WJ, Hall CB, Speers DM. Respiratory syncytial virus infection in adults: clinical, virologic, and serial pulmonary function studies. Ann Intern Med. 1978 Feb;88(2):203-5.

[123] Fine PE. The interval between successive cases of an infectious disease. Am J Epidemiol. 2003 Dec 1;158(11):1039-47.

[124] Hall CB, Geiman JM, Douglas RG, Jr., Meagher MP. Control of nosocomial respiratory syncytial viral infections. Pediatrics. 1978 Nov;62(5):728-32.

[125] Casiano-Colon AE, Hulbert BB, Mayer TK, Walsh EE, Falsey AR. Lack of sensitivity of rapid antigen tests for the diagnosis of respiratory syncytial virus infection in adults. J Clin Virol. 2003 Oct;28(2):169-74.

[126] Suwanjutha S, Sunakorn P, Chantarojanasiri T, Siritantikorn S, Nawanoparatkul S, Rattanadilok Na Bhuket T, et al. Respiratory syncytial virus-associated lower respiratory tract infection in under-5-year-old children in a rural community of central Thailand, a population-based study. J Med Assoc Thai. 2002 Nov;85 Suppl 4:S1111-9.

[127] Tupasi TE, de Leon LE, Lupisan S, Torres CU, Leonor ZA, Sunico ES, et al. Patterns of acute respiratory tract infection in children: a longitudinal study in a depressed community in Metro Manila. Rev Infect Dis. 1990 Nov-Dec;12 Suppl 8:S940-9.

[128] Berglund. Studies on respiratory syncytial virus infection. Acta Paediatric. 1967;176.

[129] Resch B, Gusenleitner W, Mandl C, Muller W. Epidemiology of respiratory syncytial virus infection in Southern Austria. Pediatr Infect Dis J. 2000 Jun;19(6):587-8.

[130] Sims DG, Downham MA, McQuillin J, Gardner PS. Respiratory syncytial virus infection in north-east England. Br Med J. 1976;2(6044):1095-8.

[131] Weigl JA, Puppe W, Schmitt HJ. Incidence of respiratory syncytial virus-positive hospitalizations in Germany. Eur J Clin Microbiol Infect Dis. 2001 Jul;20(7):452-9.

[132] Orstavik I, Carlsen KH, Halvorsen K. Respiratory syncytial virus infections in Oslo 1972--1978. I. Virological and epidemiological studies. Acta Paediatr Scand. 1980 Nov;69(6):717-22.

[133] Brandenburg AH, Jeannet PY, Steensel-Moll HA, Ott A, Rothbarth PH, Wunderli W, et al. Local variability in respiratory syncytial virus disease severity. Arch Dis Child. 1997 Nov;77(5):410-4.

[134] Cooney MK, Fox JP, Hall CE. The Seattle Virus Watch. VI. Observations of infections with and illness due to parainfluenza, mumps and respiratory syncytial viruses and Mycoplasma pneumoniae. Am J Epidemiol. 1975;101(6):532-51.

[135] Falsey AR, Walsh EE. Respiratory syncytial virus infection in elderly adults. Drugs Aging. 2005;22(7):577-87.

[136] Falsey AR, Hennessey PA, Formica MA, Cox C, Walsh EE. Respiratory syncytial virus infection in elderly and high-risk adults. N Engl J Med. 2005 Apr 28;352(17):1749-59.

[137] Kapikian AZ, Bell JA, Mastrota FM, Johnson KM, Huebner RJ, Chanock RM. An outbreak of febrile illness and pneumonia associated with respiratory syncytial virus infection. Am J Hyg. 1961;74:234-48.

[138] Monto AS, Bryan ER, Rhodes LM. The Tecumseh study of respiratory illness. VII. Further observations on the occurrence of respiratory syncytial virus and Mycoplasma pneumoniae infections. Am J Epidemiol. 1975 Dec;100(6):458-68.

[139] Madhi SA, Ismail K, O'Reilly C, Cutland C. Importance of nosocomial respiratory syncytial virus infections in an African setting. Trop Med Int Health. 2004 Apr;9(4):491-8.

[140] Doggett. Serological studies with respiratory syncytial virus. ArchGesVirusforsch. 1965;15:601-8.

[141] Hambling. A survey of antibodies to respiratory syncytial virus in the population. Br Med J. 1964;1:1223-5.

[142] Zambon M, JD Stockton, JP Clewley and DM Fleming. Contribution of influenza and respiratory syncytial virus to community cases of influenza-like illness: an observational study. Lancet. 2001;358:1410-16.

[143] Fox JP EL, Spigland I. et al. The virus watch program: A continuing surveillance of viral infections in metropolitan New York families. I Overall plan method of collecting and

handling information and a summary report of specimens collected and illnesses observed. Journal of Epidemiology. 1966;83:389-412.

[144] Gardner G, Frank AL, Taber LH. Effects of social and family factors on viral respiratory infection and illness in the first year of life. J Epidemiol Community Health. 1984 Mar;38(1):42-8.

[145] Badger G, Dingle JH, Feller AE et al. A study of illness in agroup of Cleveland families II. Incidence of common repiratory diseases. American Journal of Hygiene. 1953;58:31-40.

[146] Brimblecombe FS, Cruickshank R, Masters PL, Reid DD, Stewart GT. Family studies of respiratory infections. Br Med J. 1958 Jan 18;29(5063):119-28.

[147] Dingle JH, Badger GF, Feller AE, Hodges RG, Jordan WS, Jr., Rammelkamp CH, Jr. A study of illness in a group of Cleveland families. I. Plan of study and certain general observations. Am J Hyg. 1953 Jul;58(1):16-30.

[148] Fox JP, Hall CE, Cooney MK, Luce RE, Kronmal RA. The Seattle virus watch. II. Objectives, study population and its observation, data processing and summary of illnesses. Am J Epidemiol. 1972 Oct;96(4):270-85.

[149] Freij LaWS. Exploring child health and its ecology: The Kirkos Study in Addis Ababa. Acta Pediatr Scand. 1977(Suppl. 267).

[150] Kamath KR FR, Sundar Rao PSS and Webb JKG. Infection and Diseases in a Group of South Indian Families. II General Morbidity patterns in Families and Family Members. Am J Epidemiol. 1969(89):375-83.

[151] Feldman RA KK, Sundar Rao PSS, and Webb JKG. Infection and disease in a group of south Indian families: I Introduction, methods, definitions and general observations in a continuing study. Am J Epidemiol. 1969(89):364-74.

[152] Black RE BK, Becker S. and Yunus M. Longitudinal studies of infectious diseases and physical growth of children in rural Bangladesh. I patterns of Morbidity. Am J Epidemiol. 1982(115):305-14.

[153] ARI IWGo. Acute respiratory infections: a high priority. Tubercle and Lung Disease. 1993;74:1.

[154] ARI IWGo. Acute respiratory infections: conclusions of an IUATLD workshop. Tubercle and Lung Disease. 1993;74:2-5.

[155] Victora CG, Smith PG, Barros FC, Vaughan JP, Fuchs SC. Risk factors for deaths due to respiratory infections among Brazilian infants. Int J Epidemiol. 1989 Dec;18(4):918-25.

[156] de Francisco A, Morris J, Hall AJ, Armstrong Schellenberg JR, Greenwood BM. Risk factors for mortality from acute lower respiratory tract infections in young Gambian children. Int J Epidemiol. 1993 Dec;22(6):1174-82.

[157] Ballard TJ, Neumann CG. The effects of malnutrition, parental literacy and household crowding on acute lower respiratory infections in young Kenyan children. J Trop Pediatr. 1995 Feb;41(1):8-13.

[158] Victora CG, Huttly SR, Barros FC, Lombardi C, Vaughan JP. Maternal education in relation to early and late child health outcomes: findings from a Brazilian cohort study. Soc Sci Med. 1992 Apr;34(8):899-905.

[159] Vathanophas K, Sangchai, R., Suwanee R., Pariyanonda A., Thangsuvan J., Banyaratabhandu P., Athipanyankom S., Subharee S., Jayanetra P., Wasi C., Vorachit M. and Putnarathana P. A Community-based study of acute respiratory infection in Thai children. Rev Infect Dis. 1990;12(5):957-65.

[160] Stansfield SaS, DS. Acute respiratory infection. New York: Oxford university Press 1993.

[161] Graham N. The epidemiology of acute respiratory infections in children and adults: a global perspective. Epidemiol Rev. 1990;12:149-78.

[162] Cardoso MR, Cousens SN, de Goes Siqueira LF, Alves FM, D'Angelo LA. Crowding: risk factor or protective factor for lower respiratory disease in young children? BMC Public Health. 2004 Jun 3;4:19.

[163] Cerqueiro MC, Murtagh P, Halac A, Avila M, Weissenbacher M. Epidemiologic risk factors for children with acute lower respiratory tract infection in Buenos Aires, Argentina: a matched case-control study. Rev Infect Dis. 1990 Nov-Dec;12 Suppl 8:S1021-8.

[164] Campbell H AJaBP. Indoor air pollution in developing countries and acute respiratory infection in children [letter]. Lancet 1989;1:1012.

[165] Jim CaR, AM. Effects of passive smoking on respiratory illness from birth to age eighteen months, in Shangai, People's Republic of China. J Paediatr. 1993;123(4):553-8.

[166] Pandey MR, Boleij JS, Smith KR, Wafula EM. Indoor air pollution in developing countries and acute respiratory infection in children. Lancet. 1989 Feb 25;1(8635):427-9.

[167] Sikolia DN, Mwololo, K., Cherop, H., Hussein, A., Juma, M.,Kurui, J., Bwika, A., Osaki, Y. The Prevalence of acute respiratory infections and the associated risk factors: A Study of children under five years of age in Kibera Lindi Village, Nairobi, Kenya. J Natl Inst Public Health. 2002;51(1):67-72.

[168] Tupasi T, Velmonte,MA, Sanvictores,MEG. Determinants of morbidity and mortality due to acute respiratory infections: Implications for Infection. J Infect Dis. 1988;157:615-23.

[169] Cruz JR PGdFAPFCPaCF. Epidemiology of acute respiratory tract infections among Guatemalan ambulatory preschool children. Rev Infect Dis. 1990;12:S1029-34.

[170] Shah N, Ramankutty V, Premila PG, Sathy N. Risk factors for severe pneumonia in children in south Kerala: a hospital-based case-control study. J Trop Pediatr. 1994 Aug;40(4):201-6.

[171] Victora CG, Fuchs SC, Flores JA, Fonseca W, Kirkwood B. Risk factors for pneumonia among children in a Brazilian metropolitan area. Pediatrics. 1994 Jun;93(6 Pt 1):977-85.

[172] Johnson AW, Aderele WI. The association of household pollutants and socio-economic risk factors with the short-term outcome of acute lower respiratory infections in hospitalized pre-school Nigerian children. Ann Trop Paediatr. 1992;12(4):421-32.

[173] Dekoning HW SK, and Last JM,. Biomass Fuel Combustion and Health. Bull WHO. 1985;63(1):11-26.

[174] James J. Longitudinal Study of the morbidity of diarrheal and respiratory infections in malnourished children. Am J Clin Nutr. 1972;25:690-94.

[175] Broor S, Pandey RM, Ghosh M, Maitreyi RS, Lodha R, Singhal T, et al. Risk factors for severe acute lower respiratory tract infection in under-five children. Indian Pediatr. 2001 Dec;38(12):1361-9.

[176] Tupasi TE, Mangubat NV, Sunico MES, Magdangal DM, Navarro EE, Leonor ZA, et al. Malnutrition and acute respiratory tract infections in Filipino children. Rev Infect Dis 1990. 1990;12:S1047-54.

[177] Spooner V BJ, Tulloch S, Lehmann D, Marshall TF, Kajoi M and Alpers MP. Clinical signs and risk factors associated with pneumonia in children admitted T. Goroka Hospital, Papua New Guinea. . J Trop Pediatr. 1989 Dec;35(6):295-300.

[178] Bauchner H. LJ, Shapiro E. Studies of breast feeding and infection:how good is the evidence? JAMA. 1986;256:887-92.

[179] Fonseca W, Kirkwood BR, Victora CG, Fuchs SR, Flores JA, Misago C. Risk factors for childhood pneumonia among the urban poor in Fortaleza, Brazil: a case--control study. Bull World Health Organ. 1996;74(2):199-208.

[180] Ford K aLM. Breast feeding and child health in the United States. . J Biosoc Sci. 1993;25(2):187-94.

[181] Victora CG, Barros, F.C., Vaughan, J.P. and Reixeira, A.M.B. Birthweight and infant mortality: a longitudinal study of 5914 Brazilian children. Int J Epidemiol. 1987;16:239-45.

[182] Fenn B, Morris SS, Black RE. Comorbidity in childhood in northern Ghana: magnitude, associated factors, and impact on mortality. Int J Epidemiol. 2005 Apr;34(2):368-75.

[183] Redd SC, Bloland PB, Kazembe PN, Patrick E, Tembenu R, Campbell CC. Usefulness of clinical case-definitions in guiding therapy for African children with malaria or pneumonia. Lancet. 1992 Nov 7;340(8828):1140-3.

[184] Greenwood BM, Byass P, Greenwood AM, Hayes RJ, Menon A, Shenton FC, et al. Lack of an association between acute gastroenteritis, acute respiratory infections and malaria in young Gambian children. Trans R Soc Trop Med Hyg. 1989 Sep-Oct;83(5):595-8.

[185] Kaneko M, Watanabe J, Ueno E, Hida M, Sone T. Risk factors for severe respiratory syncytial virus-associated lower respiratory tract infection in children. Pediatr Int. 2001;43(5):489-92.

[186] Nielsen HE, Siersma V, Andersen S, Gahrn-Hansen B, Mordhorst CH, Norgaard-Pedersen B, et al. Respiratory syncytial virus infection--risk factors for hospital admission: a case-control study. Acta Paediatr. 2003 Nov;92(11):1314-21.

[187] Meissner HC. Selected populations at increased risk from respiratory syncytial virus infection. Pediatr Infect Dis J. 2003 Feb;22(2 Suppl):S40-4; discussion S4-5.

[188] Chan P, Goh A. Respiratory syncytial virus infection in young Malaysian children. Singapore Med J. 1999 May;40(5):336-40.

[189] Simoes EA. Environmental and demographic risk factors for respiratory syncytial virus lower respiratory tract disease. J Pediatr. 2003 Nov;143(5 Suppl):S118-26.

[190] Carbonell-Estrany X, Quero J, Bustos G, Cotero A, Domenech E, Figueras-Aloy J, et al. Rehospitalization because of respiratory syncytial virus infection in premature infants

younger than 33 weeks of gestation: a prospective study. IRIS Study Group. Pediatr Infect Dis J. 2000 Jul;19(7):592-7.

[191] Bulkow LR, Singleton RJ, Karron RA, Harrison LH. Risk factors for severe respiratory syncytial virus infection among Alaska native children. Pediatrics. 2002 Feb;109(2):210-6.

[192] Lanari M, Giovannini M, Giuffre L, Marini A, Rondini G, Rossi GA, et al. Prevalence of respiratory syncytial virus infection in Italian infants hospitalized for acute lower respiratory tract infections, and association between respiratory syncytial virus infection risk factors and disease severity. Pediatr Pulmonol. 2002;33(6):458-65.

[193] Wang EE, Law BJ, Stephens D. Pediatric Investigators Collaborative Network on Infections in Canada (PICNIC) prospective study of risk factors and outcomes in patients hospitalized with respiratory syncytial viral lower respiratory tract infection. J Pediatr. 1995 Feb;126(2):212-9.

[194] Weber MW, Milligan P, Hilton S, Lahai G, Whittle H, Mulholland EK, et al. Risk factors for severe respiratory syncytial virus infection leading to hospital admission in children in the Western Region of The Gambia. Int J Epidemiol. 1999;28(1):157-62.

[195] McConnochie KM, Roghmann KJ. Parental smoking, presence of older siblings, and family history of asthma increase risk of bronchiolitis. Am J Dis Child. 1986;140(8):806-12.

[196] Carbonell-Estrany X, Quero J. Hospitalization rates for respiratory syncytial virus infection in premature infants born during two consecutive seasons. Pediatr Infect Dis J. 2001 Sep;20(9):874-9.

[197] Eriksson M, Bennet R, Rotzen-Ostlund M, von Sydow M, Wirgart BZ. Population-based rates of severe respiratory syncytial virus infection in children with and without risk factors, and outcome in a tertiary care setting. Acta Paediatr. 2002;91(5):593-8.

[198] Hayes EB, Hurwitz ES, Schonberger LB, Anderson LJ. Respiratory syncytial virus outbreak on American Samoa. Evaluation of risk factors. Am J Dis Child. 1989 Mar;143(3):316-21.

[199] Roca A, F Abacassamo, M-P Loscertales, L Quintó, X Gómez-Olivé, F Fenwick, JC Saiz, G Toms and PL Alonso. Prevalence of respiratory syncytial virus IgG antibodies in infants living in rural area of Mozambique. J Med Virol. 2002;67:616-23.

[200] Adegbola RA, Falade AG, Sam BE, Aidoo M, Baldeh I, Hazlett D, et al. The etiology of pneumonia in malnourished and well-nourished Gambian children. Pediatr Infect Dis J. 1994 Nov;13(11):975-82.

[201] Hussey GD, Apolles P, Arendse Z, Yeates J, Robertson A, Swingler G, et al. Respiratory syncytial virus infection in children hospitalised with acute lower respiratory tract infection. S Afr Med J. 2000 May;90(5):509-12.

[202] Nwankwo MU, Okuonghae HO, Currier G, Schuit KE. Respiratory syncytial virus infections in malnourished children. Ann Trop Paediatr. 1994;14(2):125-30.

[203] Rooth IB, Bjorkman A. Suppression of Plasmodium falciparum infections during concomitant measles or influenza but not during pertussis. Am J Trop Med Hyg. 1992 Nov;47(5):675-81.

[204] Okoko BJ, Wesumperuma LH, Ota MOC, Pinder M, Banya W, Gomez SF, et al. The influence of placental malaria infection and maternal hypergammaglobulinemia on transplacental transfer of antibodies and IgG subclasses in a rural West African population. Journal of Infectious Diseases. 2001 Sep 1;184(5):627-32.

[205] Loscertales MP, A Roca, PJ Ventura, F Abacassamo, F dos Santos, M Sitaube, C Menendez, BM Greenwood, JC Saiz and PL Alonso. Epidemiology and clinical presentation of respiratory syncytial virus infection in a rural area of southern Mozambique. Pediatr Infect Dis J. 2002;21:148-55.

[206] Madhi SA, Venter M, Madhi A, Petersen MK, Klugman KP. Differing manifestations of respiratory syncytial virus-associated severe lower respiratory tract infections in human immunodeficiency virus type 1-infected and uninfected children. Pediatr Infect Dis J. 2001 Feb;20(2):164-70.

[207] Semple MG, Cowell A, Dove W, Greensill J, McNamara PS, Halfhide C, et al. Dual infection of infants by human metapneumovirus and human respiratory syncytial virus is strongly associated with severe bronchiolitis. J Infect Dis. 2005 Feb 1;191(3):382-6.

[208] Ruutu P, Halonen P, Meurman O, Torres C, Paladin F, Yamaoka K, et al. Viral lower respiratory tract infections in Filipino children. J Infect Dis. 1990 Feb;161(2):175-9.

[209] Miranda-Novales G, Solorzano-Santos F, Leanos-Miranda B, Vazquez-Rosales G, Palafox-Torres M, Guiscafre-Gallardo H. Blood culture and respiratory syncytial virus identification in acute lower respiratory tract infection. Indian J Pediatr. 1999 Nov-Dec;66(6):831-6.

[210] Morris M, Cohen B, Andrews N, Brown D. Stability of total and rubella-specific IgG in oral fluid samples: the effect of time and temperature. J Immunol Methods. 2002;266(1-2):111-6.

[211] Morris MC, Edmunds WJ, Miller E, Brown DW. Oral fluid collection by post--a pilot study of two approaches. Public Health. 2002;116(2):113-9.

[212] Nokes DJ, Nigatu W, Abebe A, Messele T, Dejene A, Enquselassie F, et al. A comparison of oral fluid and serum for the detection of rubella- specific antibodies in a community study in Addis Ababa, Ethiopia. Trop Med Int Health. 1998;3(4):258-67.

[213] Nokes DJ, Enquselassie F, Vyse A, Nigatu W, Cutts FT, Brown DW. An evaluation of oral-fluid collection devices for the determination of rubella antibody status in a rural Ethiopian community. Trans R Soc Trop Med Hyg. 1998;92(6):679-85.

[214] Ramsay M BR, Brown DWG. Surveillance of measles in England and Wales: implications of a national saliva testing programme. Bull WHO. 1997;75:515-21.

[215] Parry JV. Simple and reliable salivary tests for HIV and hepatitis A and B virus diagnosis asurveillance. Annals of the New York Academy of Sciences. 1993;694:216-33.

[216] Parry JV, Perry KR, Mortimer PP. Sensitive assays for viral antibodies in saliva: an alternative to tests on serum. The Lancet. 1987;2(8550):72-5.

[217] Parry JV, Perry KR, Panday S, Mortimer PP. Diagnosis of hepatitis A and B by testing saliva. Journal of Medical Virology. 1989;28(4):255-60.

[218] Nokes DJ, Enquselassie F, Nigatu W, Vyse AJ, Cohen BJ, Brown DW, et al. Has oral fluid the potential to replace serum for the evaluation of population immunity levels? A study of measles, rubella and hepatitis B in rural Ethiopia. Bull World Health Organ. 2001;79(7):588-95.

[219] Stensballe LG, Kofoed PE, Nante EJ, Sambo M, Jensen IP, Aaby P. Duration of secretory IgM and IgA antibodies to respiratory syncytial virus in a community study in Guinea-Bissau. Acta Paediatr. 2000 Apr;89(4):421-6.

[220] Toms GL, Webb MS, Milner PD, Milner AD, Routledge EG, Scott R, et al. IgG and IgM antibodies to viral glycoproteins in respiratory syncytial virus infections of graded severity. Arch Dis Child. 1989 Dec;64(12):1661-5.

[221] McIntosh K, Masters H, Orr I, Chao R, Barkin R. The immunologic response to infection with respiratory syncytial virus in infants. Journal of Infectious Diseases. 1978;138:24-32.

[222] Murphy BR, Alling DW, Snyder MH, Walsh EE, Prince GA, Chanock RM, et al. Effect of age and preexisting antibody on serum antibody response of infants and children to the F and G glycoproteins during respiratory syncytial virus infection. J Clin Microbiol. 1986;24(5):894-8.

[223] Murphy BR, Graham BS, Prince GA, Walsh EE, Chanock RM, Karzon DT, et al. Serum and nasal-wash immunoglobulin G and A antibody response of infants and children to respiratory syncytial virus F and G glycoproteins following primary infection. J Clin Microbiol. 1986;23(6):1009-14.

[224] Holmstrom P SS, Laine P, Valle SL, Suni J. HIV antibodies in whole saliva detected by ELISA and Western Blot assays. J Med Virol 1990;30:245-8.

[225] Mortimer P PJ. Non-invasive virological diagnosis: are saliva and urine specimens adequate substitutes for blood? Rev Med Virol 1991;1:73-8. 1991;1:73-8.

[226] Hazlett D, Bell, T, Tulce, P, Ademba G. Viral etiology and epidemiology of acute respiratory infections in children in Nairobi. A J of trop Med & Hyg. 1988;39:632-40.

[227] World Health Organization. Pocket book of hospital care for children: guidelines for the management of
common illnesses with limited resources.: World Health Organization 2005 2005.

[228] Meissner HC, Groothuis JR. Immunoprophylaxis and the control of respiratory syncytial virus disease. Pediatrics. 1997 Aug;100(2 Pt 1):260-3.

[229] Republic of Kenya. 1999 Population and Housing Census 2001.

[230] Government of Kenya. Kilifi District Development Plan 2002-2008: Effective Management for Sustainable Economic Growth and Poverty Reduction. In: Development MoPaN, ed.: Government of Kenya 2002:1-102.

[231] Central Bureau of Statistics. Geographic Dimensions of Well-Being in Kenya In: Departments CBoSaHRaSS, ed. *Ministry of Planning and National Development* 2003.

[232] Republic of Kenya. Kilifi District Development Plan 2002-2008.

[233] Ndiritu M, Cowgill KD, Ismail A, Chiphatsi S, Kamau T, Fegan G, et al. Immunization coverage and risk factors for failure to immunize within the Expanded Programme on Immunization in Kenya after introduction of new Haemophilus influenzae type b and hepatitis b virus antigens. BMC Public Health. 2006;6:132.

[234] Mwangi TW, Ross A, Snow RW, Marsh K. Case definitions of clinical malaria under different transmission conditions in Kilifi District, Kenya. J Infect Dis. 2005 Jun 1;191(11):1932-9.

[235] Cane P, Pringle C. Molecular epidemiology of human respiratory syncytial virus. Seminars in Virology. 1995;6:371-8.

[236] Cane PA, D. A. Matthews, and C. R. Pringle. Identification of variable domains of the attachment (G) protein of subgroup A respiratory syncytial viruses. J Gen Virol. 1991;72:2091–6.

[237] Melero JA G-BB, Martinez I, Pringle CR, Cane PA. Antigenic structure, evolution and immunobiology of human respiratory syncytial virus attachment (G) protein. J Gen Virol. 1997 Oct;78(Pt 10):2411-8.

[238] Hall CB, Douglas RG, Jr. Clinically useful method for the isolation of respiratory syncytial virus. J Infect Dis. 1975 Jan;131(1):1-5.

[239] Stockton J, JS Ellis, M Saville, JP Clewley and MC Zambon. Multiplex PCR for typing and subtyping Infuenza and Respiratory Syncytial viruses. J Clin Microbiol. 1998;36:2990-95.

[240] Cane P, CR Pringle. Molecular epidemiology of respiratory syncytial virus: rapid identification of subgroup A lineages. J Virol Methods. 1992;40:297-306.

[241] Falsey AR, McCann RM, Hall WJ, Criddle MM. Evaluation of four methods for the diagnosis of respiratory syncytial virus infection in older adults. J Am Geriatr Soc. 1996 Jan;44(1):71-3.

[242] Weigl JA, Puppe W, Grondahl B, Schmitt HJ. Epidemiological investigation of nine respiratory pathogens in hospitalized children in Germany using multiplex reverse-transcriptase polymerase chain reaction. Eur J Clin Microbiol Infect Dis. 2000 May;19(5):336-43.

[243] Ngama MJ, Ouma B, English ME, Nokes DJ. Comparison of three methods of collecting nasal specimens for respiratory virus analysis. East Afr Med J. 2004 Jun;81(6):313-7.

[244] Masters HB, Weber KO, Groothuis JR, Wren CG, Lauer BA. Comparison of nasopharyngeal washings and swab specimens for diagnosis of respiratory syncytial virus by EIA, FAT, and cell culture. Diagn Microbiol Infect Dis. 1987 Oct;8(2):101-5.

[245] Treuhaft MW, Soukup JM, Sullivan BJ. Practical recommendations for the detection of pediatric respiratory syncytial virus infections. J Clin Microbiol. 1985 Aug;22(2):270-3.

[246] Taber LH, Paredes A, Glezen WP, Couch RB. Infection with influenza A/Victoria virus in Houston families, 1976. J Hyg (Lond). 1981 Jun;86(3):303-13.

[247] Gardner PS, McQuillin J, McGuckin R. The late detection of respiratory syncytial virus in cells of respiratory tract by immunofluorescence. J Hyg (Lond). 1970 Dec;68(4):575-80.

[248] Hughes JH, Mann DR, Hamparian VV. Detection of respiratory syncytial virus in clinical specimens by viral culture, direct and indirect immunofluorescence, and enzyme immunoassay. J Clin Microbiol. 1988;26(3):588-91.

[249] John M, Cherian T, Christuraj S, John TJ. Comparison of immunofluorescence & culture for the diagnosis of respiratory syncytial virus infection. Indian J Med Res. 1990 Jul;91:242-4.

[250] Lauer BA. Comparison of virus culturing and immunofluorescence for rapid detection of respiratory syncytial virus in nasopharyngeal secretions: sensitivity and specificity. J Clin Microbiol. 1982 Aug;16(2):411-2.

[251] Reina J, Ros MJ, Del Valle JM, Blanco I, Munar M. Evaluation of direct immunofluorescence, dot-blot enzyme immunoassay, and shell-vial culture for detection of respiratory syncytial virus in patients with bronchiolitis. Eur J Clin Microbiol Infect Dis. 1995 Nov;14(11):1018-20.

[252] English M, Murphy S, Mwangi I, Crawley J, Peshu N, Marsh K. Interobserver variation in respiratory signs of severe malaria. Arch Dis Child. 1995 Apr;72(4):334-6.

[253] English M, New L, Peshu N, Marsh K. Video assessment of simple respiratory signs. Bmj. 1996 Dec 14;313(7071):1527-8.

[254] Lee M-S. An investigation of measles elimination in Taiwan: seroepidemiology and modelling. . Oxford: Oxford University; 1999.

[255] Smith P, Morrow R. Field Trials of Health Interventions in Developing Countries: a tool box. 2nd ed. London And Basingstoke: Macmillan Education Ltd. 1996.

[256] Morris MC. The application of novel oral fluid technology to age-specific surveys of immunity [Traditional Thesis]. London: University of London; 2003.

[257] Stensballe L, Poulsen A, Nante E, Jensen IP, Kofoed PE, Jensen H, et al. Mothers may transmit RSV infection more easily or severely to sons than daughters: community study from Guinea-Bissau. Scand J Infect Dis. 2004;36(4):291-5.

[258] Badger GF, Dingle JH, Feller AE, Hodges RG, Jordan WS, Jr., Rammelkamp CH, Jr. A study of illness in a group of Cleveland families. V. Introductions and secondary attack rates as indices of exposure to common respiratory diseases in the community. Am J Hyg. 1953 Sep;58(2):179-82.

[259] Berman S. Epidemiology of acute respiratory infections in children of developing countries. Rev Infect Dis. 1991 May-Jun;13 Suppl 6:S454-62.

[260] Hurrell GD, Sturdy PM, Frood JD, Gardner PS. Viruses in families. Lancet. 1971 Apr 17;1(7703):769-74.

[261] Falsey AR, Criddle MC, Walsh EE. Detection of respiratory syncytial virus and human metapneumovirus by reverse transcription polymerase chain reaction in adults with and without respiratory illness. J Clin Virol. 2006 Jan;35(1):46-50.

[262] Albargish KA, Hasony HJ. Respiratory syncytial virus infection among young children with acute respiratory tract infection in Iraq. East Mediterr Health J. 1999 Sep;5(5):941-8.

[263] Stansfield SK. Acute respiratory infections in the developing world: strategies for prevention, treatment and control. Pediatr Infect Dis J. 1987 Jul;6(7):622-9.

[264] van Ginneken JK, Lob-Levyt J, Gove S. Potential interventions for preventing pneumonia among young children in developing countries: promoting maternal education. Trop Med Int Health. 1996 Jun;1(3):283-94.

[265] Lim YW, Steinhoff M, Girosi F, Holtzman D, Campbell H, Boer R, et al. Reducing the global burden of acute lower respiratory infections in children: the contribution of new diagnostics. Nature. 2006 Nov 23;444 Suppl 1:9-18.

[266] Zar HJ, Mulholland K. Global burden of pediatric respiratory illness and the implications for management and prevention. Pediatr Pulmonol. 2003 Dec;36(6):457-61.

[267] Abebe A, Nokes DJ, Dejene A, Enquselassie F, Messele T, Cutts FT. Seroepidemiology of hepatitis B virus in Addis Ababa, Ethiopia: transmission patterns and vaccine control. Epidemiol Infect. 2003 Aug;131(1):757-70.

[268] Cutts FT, Abebe A, Messele T, Dejene A, Enquselassie F, Nigatu W, et al. Sero-epidemiology of rubella in the urban population of Addis Ababa, Ethiopia. Epidemiol Infect. 2000 Jun;124(3):467-79.

[269] Enquselassie F, Ayele W, Dejene A, Messele T, Abebe A, Cutts FT, et al. Seroepidemiology of measles in Addis Ababa, Ethiopia: implications for control through vaccination. Epidemiol Infect. 2003 Jun;130(3):507-19.

[270] Filmer D, Pritchett LH. Estimating wealth effects without expenditure data--or tears: an application to educational enrollments in states of India. Demography. 2001 Feb;38(1):115-32.

[271] StataCorp. Stata reference manual H-P release 8 Texas: Stata Corporation 2003.

[272] StataCorp. Stata Release 9: Multivariate Statistics Texas: StataCorp 2005.

[273] WHO. WHO Child Growth Standards STATA Macro (igrowup.ado). World Health Organization 2006.

[274] Multicentre Growth Reference Study Group. WHO Child Growth Standards: Length/height-for-age, weight-for-age, weight-for-length, weight-for-height and body mass index-for-age: Methods and development. Geneva World Health Organization; 2006.

[275] David Clayton, Hills M. Statistical Models in Epidemiology. Oxford: Oxford University Press 2001.

[276] StataCorp. NetCourse 631: an Introduction to survival analysis. 2000 [cited; Available from: http://www.stata.com.nc631/lectures/nc631

[277] Cleves M, Gould W, Gutierrez R. An Introduction to Survival Analysis Using Stata 2004.

[278] Stata reference manual release 8. :Survival Analysis and Epidemiological Tables. Texas: Stata Corporation 2003.

[279] Gutierrez R. st: how to use stcox in time-varying covariate model. 2006 [cited 2006 31 May 2006]; Available from:

[280] UCLA. Stata FAQ: Testing teh proportional hazard assumption in Cox models. [cited 2006 19/05/2006]; Available from:
http://www.ats.ucla.edu/stat/stata/faq/test_proportionality.htm

[281] Gutierrez R. st: Goodness of fit using Cox-snell residuals. 2006 [cited 2006 01 June 2006]; Available from:

[282] UCLA. Introduction to survival Analysis with Stata Seminar. [cited 2005 26/10/2005]; Available from:
http://www.ats.ucla.edu/STAT/stata/seminars/stata_survival/default.htm

[283] Dahoo I, Martin W, Stryhn H Veterinary Epidemiologic Research. First ed. Prince Edward Island: AVC Inc 2003.

[284] World Bank. Measuring Living Standards:Household Consumption and Wealth Indices: World Bank.

[285] EQUINET. Equity in Health Care in Namibia Towards needs-based allocation formula. In: Ministry of Health and Social Services N, ed.: Ministry of Health and Social Services, Namibia World Health Organisation, Regional Network for Equity in Health in Southern Africa 2005.

[286] Berkley J, Mwangi I, Griffiths K, Ahmed I, Mithwani S, English M, et al. Assessment of severe malnutrition among hospitalized children in rural Kenya: comparison of weight for height and mid upper arm circumference. Jama. 2005 Aug 3;294(5):591-7.

[287] Vardas E, Blaauw D, McAnerney J. The epidemiology of respiratory syncytial virus (RSV) infections in South African children. S Afr Med J. 1999;89(10):1079-84.

[288] Chandra RK. Nutrition, immunity, and infection: present knowledge and future directions. Lancet. 1983 Mar 26;1(8326 Pt 1):688-91.

[289] Neumann CG, Lawlor GJ, Jr., Stiehm ER, Swenseid ME, Newton C, Herbert J, et al. Immunologic responses in malnourished children. Am J Clin Nutr. 1975 Feb;28(2):89-104.

[290] Reddy V, Bhaskaram C, Raghuramulu N. Immunological responses in malnourished children. Indian Pediatr. 1977 Apr;14(4):255-8.

[291] Watson RR, McMurray DN, Martin P, Reyes MA. Effect of age, malnutrition and renutrition on free secretory component and IgA in secretions. Am J Clin Nutr. 1985 Aug;42(2):281-8.

[292] Schlesinger L, Arevalo M, Arredondo S, Diaz M, Lonnerdal B, Stekel A. Effect of a zinc-fortified formula on immunocompetence and growth of malnourished infants. Am J Clin Nutr. 1992 Sep;56(3):491-8.

[293] Ashford RW, Craig PS, Oppenheimer SJ. Polyparasitism on the Kenya coast. 1. Prevalence, and association between parasitic infections. Ann Trop Med Parasitol. 1992 Dec;86(6):671-9.

[294] DeVincenzo JP. Factors predicting childhood respiratory syncytial virus severity: what they indicate about pathogenesis. Pediatr Infect Dis J. 2005 Nov;24(11 Suppl):S177-83, discussion S82.

[295] Chuma JM, Thiede M, Molyneux CS. Rethinking the economic costs of malaria at the household level: evidence from applying a new analytical framework in rural Kenya. Malar J. 2006;5:76.

[296] Bor DH, Epstein PR. Pathogenesis of respiratory infection in the disadvantaged. Semin Respir Infect. 1991 Dec;6(4):194-203.

[297] Simiyu DE, Wafula EM, Nduati RW. Mothers' knowledge, attitudes and practices regarding acute respiratory infections in children in Baringo District, Kenya. East Afr Med J. 2003 Jun;80(6):303-7.

[298] Martinez FD, Morgan WJ, Wright AL, Holberg CJ, Taussig LM. Diminished lung function as a predisposing factor for wheezing respiratory illness in infants. N Engl J Med. 1988 Oct 27;319(17):1112-7.

[299] Hall CB, Douglas RG, Jr. Modes of transmission of respiratory syncytial virus. J Pediatr. 1981;99(1):100-3.

[300] Kim HW, Canchola JG, Brandt CD, Pyles G, Chanock RM, Jensen K, et al. Respiratory syncytial virus disease in infants despite prior administration of antigenic inactivated vaccine. Am J Epidemiol. 1969 Apr;89(4):422-34.

[301] Nokes DJ, Anderson RM. The use of mathematical models in the epidemiological study of infectious diseases and in the design of mass immunization programmes. Epidemiol Infect. 1988 Aug;101(1):1-20.

[302] Hall CB, Powell KR, MacDonald NE, Gala CL, Menegus ME, Suffin SC, et al. Respiratory syncytial viral infection in children with compromised immune function. N Engl J Med. 1986 Jul 10;315(2):77-81.

[303] Hall CB, Douglas RG, Jr., Schnabel KC, Geiman JM. Infectivity of respiratory syncytial virus by various routes of inoculation. Infect Immun. 1981 Sep;33(3):779-83.

[304] Hacimustafaoglu M, Celebi S, Aynaci E, Sinirtas M, Koksal N, Kucukerdogan A, et al. The progression of maternal RSV antibodies in the offspring. Arch Dis Child. 2004;89(1):52-3.

[305] Ward KA, Lambden PR, Ogilvie MM, Watt PJ. Antibodies to respiratory syncytial virus polypeptides and their significance in human infection. J Gen Virol. 1983 Sep;64 (Pt 9):1867-76.

Appendices

1. Field Procedures Appendices

The attached appendices give details of field and sample collection procedures used during follow up for both the main cohort and the household study. Appendix A is the questionnaire used in the MCH clinic exit poll, a tool that was used to select the study area (Chapter 3). Following this is a series of all the consent forms used during the study (Appendix B-E). The next few appendices (F & G) provide specific details on field and research Out Patient clinic study procedures including the standard operating procedures (SOPs) for sample (nasal, oral-fluid and blood) collection and the limitations of cost covered during the study.

Appendix A: Exit Poll Questionnaire

EXIT POLL QUESTIONNAIRE

Date [ ][ ][ ][ ][ ][ ][ ][ ]

Where did you give birth? [ ] Home, KDH, HC, Other

If Other specify _____

Date of birth? [ ][ ][ ][ ][ ][ ][ ][ ] (RE: Road to Health Card or RHC)

How old is the baby now? Months [ ] Weeks [ ]

How old was the baby at the first clinic visit? Months [ ] Weeks [ ]

What date was that? (RE:RHC) [ ][ ][ ][ ][ ][ ][ ][ ]

What is the reason for your visit? _____

_____ Immunisation (list if more than one) / Other (specify)

Where do you live? Name of village _____

Location _____

Sub-location _____

Bus fare from home to the hospital (one way) [ ]

How long does it take you to get to the hospital (<15mins, 15-<30mins, 30-<60mins, ≥60mins) _____

Consent Forms

The following is a listing of the consent forms that will follow

1. RSV Birth Cohort consent form - this is the consent form used to recruit cohort infants at birth into the main RSV birth cohort (Appendix B).

2. Household study consent form – this is the consent form used to recruit household members of cohort infant into the household study. In this initial household consent form only siblings < 15 years were recruited into clinical surveillance similar to that for the cohort child. This form is almost exactly the same as the RSV Birth Cohort consent form except that it is used in reference to all children in the home who were under 15 years at the time of recruitment. **Differences between consent forms are underlined.** The involvement of adults (household member 15 years or older) was limited to having 3-monthly oral-fluid samples collected (Appendix C).

3. Adult consent form - after receiving approval form the various ethical and scientific committees, adult household members were asked to consent to routine clinical surveillance as for the children in the home using this consent form. Consent form is same as that used for household children but changed to reflect reference to adult family members (Appendix D).

4. Follow of study children and households had to be extended for an additional year (explained in chapter 3). Appendix E is the consent form used to request an extension of follow of study households.

Appendix B: Information and consent form for birth cohort child.

STUDY OF A MAJOR CAUSE OF PNEUMONIA IN KILIFI DISTRICT

Information and consent form for birth cohort child.

PURPOSE OF STUDY

Chest infections are one of the biggest health problems for children in Kilifi District. We know the cause and the right treatment for many of these, but in some cases this is not yet known. We are investigating an infection caused by a germ (virus) that is known to be a major cause of chest infections throughout the world. The virus is called Respiratory Syncytial Virus or RSV. For infants who are infected in their first year of life the symptoms can be severe enough for the parents to feel a need to take the child to hospital. This condition is distressing both to the infant and to the family, although is rarely life threatening. There are no drugs in routine use for reducing the symptoms, and currently no vaccine to prevent infection. However, much work is being carried out to develop a vaccine to prevent infection.

The purpose of this study is to identify how important is RSV in Kilifi District, and to find out how the virus is transmitted and maintained within the population. Results of both would be of great value in developing and eventually implementing a vaccine against this disease.

WHO IS ORGANISING THE STUDY

The study is organised by the Kenya Medical Research Institute (KEMRI) CGMRC, Kilifi District Hospital (KDH) in collaboration with groups from outside Kenya (Universities in England) who have specific knowledge of the disease under study.

YOUR INVOLVEMENT

You are already enrolled in a Birth Cohort Study of Malaria and Pneumonia (pneumococcal) disease run by KEMRI CGMRC. We are asking you to involve your newly born child in another study of pneumonia in which we will monitor the child for infections of the nose and lungs that may be due to RSV. The method of monitoring is EITHER through you bringing your child to the hospital when he / she has symptoms of a cold or a chest infection, OR through weekly visits during the outbreaks of these problems (which occur each year) of our survey staff to your house to check on the child's health. If you decide to enroll we will record information on your pregnancy, details of your child (weight, length), and household information (location, number of siblings, household details).

We will monitor your child until the end of year 2004 which is a period of three years [insert 2 years as appropriate].

If the child has signs of nose or chest infection that might be due to RSV we will collect a nasal specimen (fluid from the nose) in which we will check for the virus. If our tests confirm that the child is infected with RSV we will request a small sample of blood (less than half a teaspoon) and saliva soon after and again after 3-4 weeks by which to see how the child's immune defenses are reacting.

Additionally, every 3 months we will collect a sample of saliva. You have already agreed for blood samples to be collected from your child every three months as part of the malaria and pneumococcal pneumonia study. For this RSV study, we would like to use a small amount of these blood samples and {where appropriate} we would also like you to continue to bring your child in for 3 monthly blood sampling until the end of the study period. The samples are to look for evidence of the infection.

{Where appropriate} We will also ask for a swab from the nose and throat each week in the first year during the main season of RSV, to assist in identifying the disease.

HOW DO WE COLLECT SAMPLES

Nasal specimens will be collected by washing out the nasal passage with some harmless fluid. The fluid is squirted into one nostril using a syringe [demonstrate device] and collected as it returns through the same or the other nostril. This procedure, conducted by trained staff, causes no harm to the infant, and whilst being a little uncomfortable, should clear the nasal passage making breathing more easy.

We collect saliva using a small clean sponge on the end of a plastic stick [show the device]. The sponge is brushed around the teeth and gums for half a minute - like using a toothbrush. This is painless, and without risk. Our staff will show how it is done.

Blood collection is by trained personnel. We will collect a sample of 2ml as much as will fit in this tube [demonstrate a 2ml tube].

{Where appropriate} Nasal and throat swabs are collected using cotton buds on the end of a wire support. This brushed against the back of the throat (throat swab [show device]) or inserted into the nose and brushed against the wall of the inner nose (nasal swab [show device]).

RISKS AND BENEFITS OF THE STUDY

All of the tests described above (taking blood or collecting fluid from the nose or mouth) are commonly used tests in clinics in many countries. They carry no danger for your child. Some of the procedures may be slightly uncomfortable, we try to ensure that this is not a problem by using highly trained staff to perform the tests.

Your child will benefit from having coughs, chest infections and fever investigated more quickly than would usually be the case because they are being seen more often. If your child is unwell at any time during the study, s/he can attend the research clinic at KDH and we will provide, without charge, examination and routine investigations (such as malaria tests) and medications (such a treatment for malaria, chest, skin or ear infections). Throughout the study period if our field worker asks you to attend the clinic with the child because s/he is unwell, or you bring the child to the clinic because you yourself have identified signs of nose or lung infection in the child (signs that we will explain), we will refund reasonable costs of your return travel fare. Because our funds are strictly limited we can only offer these benefits to the child involved in the study.

The other main benefit of research studies is in helping to improve understanding of disease, which may lead to better ways of preventing and treating these diseases. This benefits children and families in Kilifi District as a whole, as well as in other parts of Kenya.

WHAT HAPPENS TO THESE SAMPLES AND THE INFORMATION COLLECTED

The samples are sent to the laboratory at Kilifi Hospital where they will be tested to identify if your child is infected with RSV, or has had the infection in the recent past. To protect the privacy of your child, we will keep the records of this study under a code number rather than by name. We will keep the records in locked files and only study staff will be allowed to look at them. The name of your child or other facts that might point to him/her will not appear when we present this study or publish its results.

We would also like to store any blood/saliva/nasal washings that are left over after we do

your tests. We plan to use these samples for studies we may do in the future. The samples will be labeled with a number, not a name, and if further information is required, e.g. to identify that this sample comes from your child, this information will only be provided to the researchers if the new study has been scrutinised and approved by the National Ethical Committee in KEMRI. You can decline to let us store your child's blood and saliva for use in other research projects and still be in this study.

MORE INFORMATION

Please feel free to ask any questions about the study. You are free to join in the study or not. You may also decide to withdraw your child from the study at any time, for any reason. If, after discussing this with us, you decide to allow your child to join the study we would like you to sign this form.

We have given you a copy of this consent form. When you sign below, it shows that you agree to join the study. If there is any part of this form which you do not understand be sure to ask questions about it. Do not sign until you have full answers to all your questions. When you are ready to be part of the study, please tick one of the two boxes and print and sign your name on the lines below.

Consent form for infant cohort study.

STUDY OF A MAJOR CAUSE OF PNEUMONIA IN KILIFI DISTRICT

To Mother

- Have you read, or had read to you, the invitation letter?
- Have you had an opportunity to ask questions and discuss this study?
- Have you received satisfactory answers to those questions?
- Have you received enough information about the study?
- Do you understand that you are free to withdraw your child from the study
  - at any time
  - without having to give a reason why
  - without affecting your child's access to medical provision

I wish my child [name] _____to take part in the study

☐ **I also give consent for my child's blood/saliva/nasal samples to be stored at KEMRI under the conditions above**

☐ **I do not give consent for my child's blood to be stored at KEMRI for future research**

Print Name: _____

Signature _____ Date _____

Household number/registration _____

Witness

I observed the process of consent. The parent/guardian of the prospective participant read this form, was given the chance to ask questions, appeared to accept the answers, and signed to enroll his/her child in the study.

Name: _____

Signature _____ Date _____

The study co-ordinator is Dr James Nokes. If you have any questions about the ethics of this study you can contact Dr Charles Mbogo, at KEMRI/Wellcome Trust, Kilifi District Hospital or telephone Kilifi 22535.

Appendix C: Information and consent form for household study participants.

HOUSEHOLD STUDY OF A MAJOR CAUSE OF PNEUMONIA IN KILIFI DISTRICT

Information and consent form.

PURPOSE OF STUDY

Chest infections are one of the biggest health problems for children in Kilifi District. We know the cause and the right treatment for many of these, but in some cases this is not yet known. We are investigating an infection caused by a germ (virus) that is known to be a major cause of chest infections throughout the world. The virus is called Respiratory Syncytial Virus or RSV. For infants who are infected in their first year of life the symptoms can be severe enough for the parents to feel a need to take the child to hospital. This condition is distressing both to the infant and to the family, although is rarely life threatening. In older children and adults the infection is usually less severe. There are no drugs in routine use for reducing the symptoms, and currently no vaccine to prevent infection. However, much work is being carried out to develop a vaccine to prevent infection.

The purpose of this study is to identify how important is RSV in Kilifi District, and to find out how the virus is transmitted and maintained within the population. Results of both would be of great value in developing and eventually implementing a vaccine against this disease.

WHO IS ORGANISING THE STUDY

The study is organised by the Kenya Medical Research Institute (KEMRI) CGMRC, Kilifi District Hospital (KDH) in collaboration with groups from outside Kenya (Universities in England) who have specific knowledge of the disease under study.

YOUR INVOLVEMENT

Your newborn child has been recruited into a study of pneumonia in which we will monitor the child for infections of the nose and lungs that may be due to RSV. There will eventually be around 600 other newborns in this study from Kilifi District. The method of monitoring your child is EITHER through you bringing your child to the hospital when he / she has symptoms of a cold or a chest infection, OR through visits every month at least and weekly during the outbreaks of the virus (which occurs each year) of our survey staff to your house to check on the child's health. We are now asking you and your whole family to be involved in a household study of the same virus. We are asking around 70 other households in the District also. In this study the infant's siblings (your other children) who are under 15 years of age will be monitored in the same way (mentioned above) as the infant. In addition, as for your infant, we will request routine saliva samples from all household members at enrollment and every three months thereafter until the end of the study (end of 2004). If you decide to enroll your family into this study we will record information on all the members of your household and ask you to sign this consent form.

This study will continue until the end of year 2004, that is, a period of up to two years.

During this period if any of your children (those under 15 years) has signs of nose or chest infection that might be due to RSV we will collect a nasal specimen (fluid from the nose) in which we will check for the virus. If our tests confirm that the child is infected with RSV we will request a small sample of blood (less than half a teaspoon) and saliva soon after and again after 3-4 weeks by which to see how the child's immune defenses are reacting. We will also increase the number of oral-fluid samples collected over the next 3 months, with a

sample every week. These samples will assist in understanding how you develop your protection to RSV infection.

{if appropriate} From 40 of the 70 households recruited we will ask to take blood samples from the whole family on two consecutive visits (that is, one following the other) at the time of the 3 monthly saliva samples. We would like your family to be one of these families, and if you agree we would collect blood samples beginning on the {first, second, third ...as appropriate} 3 month visit. .We require these samples for comparison and confirmation of the saliva results. Before taking any such blood we will remind you about why it is required.

{households selected for intensive sampling} Finally, some of your household {this will be known by the field worker at that time} will be requested to be involved in an intensive study of oral fluid. At the time of this study we will ask for one oral-fluid sample and for the subsequent 13 weekly visits. This will assist us to understand more accurately how each person responds to RSV infection and re-infection.

HOW DO WE COLLECT SAMPLES

Nasal specimens will be collected by washing out the nasal passage with some harmless fluid. The fluid is squirted into one nostril using a syringe [demonstrate device] and collected as it returns through the same or the other nostril. This procedure, conducted by trained staff, causes no harm to the infant, and whilst being a little uncomfortable, should clear the nasal passage making breathing more easy.

We collect saliva using a small clean sponge on the end of a plastic stick [show the device]. The sponge is brushed around the teeth and gums for half a minute - like using a toothbrush. This is painless, and without risk. Our staff will show how it is done.

Blood collection is by trained personnel. We will collect a sample of 2ml as much as will fit in this tube [demonstrate a 2ml tube].

RISKS AND BENEFITS OF THE STUDY

All of the tests described above (taking blood or collecting fluid from the nose or mouth) are commonly used tests in clinics in many countries. They carry no danger for you or your children. Some of the procedures may be slightly uncomfortable; we try to ensure that this is not a problem by using highly trained staff to perform the tests.

Your children will benefit from having coughs, chest infections and fever investigated more quickly than would usually be the case because they are being seen more often. If your children are unwell at any time during the study, they can attend the research clinic at KDH and we will provide, without charge, examination and routine investigations (such as malaria tests) and medications (such as treatment for malaria, chest, skin or ear infections). Throughout the study period if our field worker asks you to attend the clinic with a child because s/he is unwell, or you bring the child to the clinic because you yourself have identified signs of nose or lung infection in the child (signs that we will explain), we will refund reasonable costs of your return travel fare. Because our funds are strictly limited we cannot afford to provide this service for anyone who is not in the study.

The other main benefit of research studies is in helping to improve understanding of disease, which may lead to better ways of preventing and treating these diseases. This benefits children and families in Kilifi District as a whole, as well as in other parts of Kenya.

WHAT HAPPENS TO THESE SAMPLES AND THE INFORMATION COLLECTED

The samples are sent to the laboratory at Kilifi Hospital where they will be tested to identify if your child(ren) is infected with RSV, or has had the infection in the recent past. To protect the privacy of your family, we will keep the records of this study under a code number rather than by name. We will keep the records in locked files and only study staff will be allowed to look at them. The names of your family members or other facts that might point to them will not appear when we present this study or publish its results.

We would also like to store any blood/saliva/nasal washings that are left over after we do your tests. We plan to use these samples for studies we may do in the future. The samples will be labeled with a number, not a name, and if further information is required, e.g. to identify that this sample comes from your child, this information will only be provided to the researchers if the new study has been scrutinised and approved by the National Ethical Committee in KEMRI. You can decline to let us store leftover blood and saliva samples for use in other research projects and still be in this study.

MORE INFORMATION

Please feel free to ask any questions about the study. You are free to join in the study or not. You may also decide to withdraw from the study at any time, for any reason. If, after discussing this with us, you decide to allow your family to join the study we would like you to sign this form.

We have given you a copy of this consent form. When you sign below, it shows that you agree to join the study. If there is any part of this form which you do not understand be sure to ask questions about it. Do not sign until you have full answers to all your questions. When you are ready to be part of the study, please tick one of the two boxes and print and sign your name on the lines below.

Consent form for household RSV study (page 1).

STUDY OF A MAJOR CAUSE OF PNEUMONIA IN KILIFI DISTRICT

To Mother/Father

- Have you read, or had read to you, the invitation letter?
- Have you had an opportunity to ask questions and discuss this study?
- Have you received satisfactory answers to those questions?
- Have you received enough information about the study?
- Do you understand that you are free to withdraw your child from the study
  - at any time
  - without having to give a reason why
  - without affecting your child's access to medical provision

	Name of participant	I agree to participate	I agree to specimen storage	Date	Witness name*	Signature	Date
1							
2							
3							
4							
5							
6							

Continue on separate page if more than 6 adults

Consent form for household RSV study (page 2).

I agree to the involvement of my children (under 15 years of age) in the study:

child/children's [name] :

1. _____ **2.** _____ **3.** _____

4. _____ **5.** _____ **6.** _____

7. _____ **8.** _____ **9.** _____

continue on separate page if more

☐ **I also give consent for blood/saliva/nasal samples collected to be stored at KEMRI under the conditions above**

☐ **I do not give consent for samples colleted to be stored at KEMRI for future research**

Mother/Father Name: _____ **signature:** _____ **Date:** _____

***Witness**

I observed the process of consent. The parent/guardian of the prospective participant read this form, was given the chance to ask questions, appeared to accept the answers, and signed to enroll his/her child in the study.

Name: _____

Signature _____ Date _____

The study co-ordinator is Emelda Okiro. If you have any questions about the ethics of this study you can contact Dr Charles Mbogo, at KEMRI/Wellcome Trust, Kilifi District Hospital or telephone Kilifi 22535.

Appendix D: Information and consent form for Adults in the Household study.

HOUSEHOLD STUDY OF A MAJOR CAUSE OF PNEUMONIA IN KILIFI DISTRICT

Information and consent form for Adults.

PURPOSE OF STUDY

As you know we are investigating an infection caused by a germ (virus) that is known to be a major cause of chest infections throughout the world. The virus is called Respiratory Syncytial Virus or RSV. For infants who are infected in their first year of life the symptoms can be severe enough for the parents to feel a need to take the child to hospital. This condition is distressing both to the infant and to the family, although is rarely life threatening. In older children and adults the infection is usually less severe. There are no drugs in routine use for reducing the symptoms, and currently no vaccine to prevent infection. However, much work is being carried out to develop a vaccine to prevent infection.

The purpose of the household study in which you are already involved is to identify who is bringing the infection into the home. Results of this would be of great value in protecting the infant in whom (as stated above) this infection is most dangerous.

YOUR ADDITIONAL INVOLVEMENT

Your whole family has been recruited into the household study of RSV. In this study your children those of whom are under 15 years of age are monitored by EITHER you bringing your children to the hospital when they have symptoms of a cold or a chest infection, OR through visits every month at least and weekly during the outbreaks of the virus (which occurs each year) of our survey staff to your house to check on your children's health. We are also taking saliva samples from all members of your family every three months. We are now inviting older members within the family to be involved in a similar monitoring system. During this period if you (or other adult family member) have signs of nose or chest infection that might be due to RSV we will collect a nasal specimen (fluid from the nose) in which we will check for the virus. If our tests confirm that you (or other family member) are infected with RSV we will request a small sample of blood (less than half a teaspoon) and saliva soon after and again after 3-4 weeks by which to see how the immune defenses are reacting. If you decide to enroll we will require you to sign this consent form.

This study will continue until the end of year 2004.

HOW DO WE COLLECT SAMPLES

Nasal specimens will be collected by washing out the nasal passage with some harmless fluid. The fluid is squirted into one nostril using a bulb syringe [demonstrate device] and collected as it returns through the same or the other nostril. This procedure, conducted by trained staff, causes no harm to the infant, and whilst being a little uncomfortable, should clear the nasal passage making breathing more easy.

RISKS AND BENEFITS OF THE STUDY

All of the tests described above (taking blood or collecting fluid from the nose or mouth) are commonly used tests in clinics in many countries. They carry no danger for you. Some of the procedures may be slightly uncomfortable; we try to ensure that this is not a problem by using highly trained staff to perform the tests.

You will benefit from having coughs, chest infections and fever investigated more quickly than would usually be the case because you are being seen more often. If you are unwell at any time during the study, you can attend the research clinic at KDH and we will provide, without charge, examination and routine investigations (such as malaria tests) and medications (such as treatment for malaria, chest, skin or ear infections). Throughout the study period if you present at the clinic because you have identified signs of nose or lung infection in the child (signs that we will explain), we will refund reasonable costs of your return travel fare. Because our funds are strictly limited we cannot afford to provide this service for anyone who is not in the study.

The other main benefit of research studies is in helping to improve understanding of disease, which may lead to better ways of preventing and treating these diseases. This benefits children and families in Kilifi District as a whole, as well as in other parts of Kenya.

WHAT HAPPENS TO THESE SAMPLES AND THE INFORMATION COLLECTED

The samples are sent to the laboratory at Kilifi Hospital where they will be tested to identify if you are infected with RSV, or have had the infection in the recent past. To protect your privacy, we will keep the records of this study under a code number rather than by name. We will keep the records in locked files and only study staff will be allowed to look at them. Your name or that of any family members or other facts that might point to them will not appear when we present this study or publish its results.

We would also like to store any specimens that are left over after we do your tests. We plan to use these samples for studies we may do in the future. The samples will be labeled with a number, not a name, and if further information is required, e.g. to identify that this sample comes from your child, this information will only be provided to the researchers if the new study has been scrutinised and approved by the National Ethical Committee in KEMRI. You can decline to let us store leftover blood and saliva samples for use in other research projects and still be in this study.

MORE INFORMATION

Please feel free to ask any questions about the study. You are free to join in the study or not. You may also decide to withdraw from the study at any time, for any reason. If, after discussing this with us, you decide to allow your family to join the study we would like you to sign this form.

We have given you a copy of this consent form. When you sign below, it shows that you agree to join the study. If there is any part of this form which you do not understand be sure to ask questions about it. Do not sign until you have full answers to all your questions. When you are ready to be part of the study, please tick one of the two boxes and print and sign your name on the lines below.

STUDY OF A MAJOR CAUSE OF PNEUMONIA IN KILIFI DISTRICT

- Have you read, or had read to you, the invitation letter?
- Have you had an opportunity to ask questions and discuss this study?
- Have you received satisfactory answers to those questions?
- Have you received enough information about the study?
- Do you understand that you are free to withdraw from the study
  - at any time
  - without having to give a reason why
  - without affecting your normal access to medical provision

Name of participant	I agree to participate	I agree to specimen storage	Date	Participant signature	Witness name*	Signature	Date

Continue on separate page if more than 6 adults

Appendix E: Information and consent form for RSV Study extension.

STUDY EXTENSION

STUDY OF A MAJOR CAUSE OF PNEUMONIA IN KILIFI DISTRICT

Why we are here
Your family (yourself and your children) are involved in a study of a germ (known as RSV) that is known to be a major cause of chest infections throughout the world. We would first like to thank you for your involvement up to now.

From our study we now know that RSV is a serious problem in Kilifi. For example, we found that 1 in 100 children in the study were admitted to Kilifi District Hospital (KDH) through RSV infection in the first year of life.

We originally asked you to participate for the first two years of your child's (the index) life. This was in order to learn more about infections in the children's first and second years of life how they get infected. However, in the first year (2002) fewer of the 300 children that we were following were infected than we expected. We therefore would like to continue to monitor children and therefore your family for another RSV season, that is expected next year (2005).

What are we asking for
We are therefore inviting you and your children to continue to participate in the study for another year, which is until October 2005 at the latest.

What changes come with continued participation
If you agree to continue to participate there is no change in the way you and your family participate or the compensation that we offer.

In summary:

- We will make house visits every week during the RSV germ season (first half of year) when it is known to cause more illness in community, and monthly otherwise. You are encouraged to take your children {or yourself} to the Out-Patient research clinic at KDH if there are signs of nose or chest infection.
- If your children {or yourself} have signs of nose or chest infection that might be due to RSV we will collect a nasal specimen (fluid from the nose) in which we will check for the virus.
- If our tests confirm that {you or} your child(ren) is infected with RSV we will request a small sample of blood (less than half a teaspoon) and saliva soon after and again after 3-4 weeks by which to see how the child's immune defences are reacting.
- Additionally, every 3 months we will collect a sample of saliva.
- If your child is participating in the pneumococcal pneumonia study (Kilifi Birth Cohort) and you have agreed to allow us to use a small amount of the blood collected every 3 months, we ask you to continue providing this sample for a third year. The blood and saliva samples are to look for evidence of infection that we have missed by our observation.

If you would like us to go over the details of your involvement, how we collect samples, what we do with them and the risk and benefits of participation then please say so.

More information
Please feel free to ask any questions about the study. You are free to join in the study or not, and may also withdraw your child from the study at any time, for any reason. If, after discussing this with us, you decide to allow your child to continue to participate in the study we would like you to sign this form.

We have given you a copy of this consent form. When you sign below, it shows that you agree to join the study. If there is any part of this form which you do not understand be sure to ask questions about it. Do not sign until you have full answers to all your questions. When you are ready to be part of the study, please tick one of the two boxes and print and sign your name on the lines below.

Consent form for household study.

STUDY OF A MAJOR CAUSE OF PNEUMONIA IN KILIFI DISTRICT

To Mother/Father

- Have you read, or had read to you, the invitation letter?
- Have you had an opportunity to ask questions and discuss this study?
- Have you received satisfactory answers to those questions?
- Have you received enough information about the study?
- Do you understand that you are free to withdraw from the study
  - at any time
  - without having to give a reason why
  - without affecting your child's access to medical provision

I wish my family to continue to participate in the study

Print Parents Name: _____

Print Parents Name: _____

Child name _____ **Child name** _____

Child name _____ **Child name** _____

Child name _____ **Child name** _____

Child name _____ **Child name** _____

Child name _____ **Child name** _____

Signature _____ Date _____

Witness
I observed the process of consent. The parent/guardian of the prospective participant was read

this form, was given the chance to ask questions, appeared to accept the answers, and signed to enroll his/her child in the study.

Name: _____

Signature _____ Date _____

The study co-ordinators are Dr James Nokes and Emelda Okiro. If you have any questions about the ethics of this study you can contact Dr Charles Mbogo, at KEMRI/Wellcome Trust, Kilifi District Hospital or telephone Kilifi 22535.

Appendix F: Active Surveillance Form

RSV FOLLOW-UP

Date: ____/____/____

Next appt : ____/____/____

KBC-No.	RSV-No	NAME1	NAME2	NAME3	SEX	FLWK	TEMP	RSV+2wk	BDWK	KEYSGNWK	BDOBS	KEYSGNOB	NWREQ	BRATE	REFER	AWAYWK
3245		JUMA	CHE	ALI	M	N	36.41	N	Y	C+DC	N	F+C	Y 10:00am	60/min	Y 30/=	2wks

Key signs: Difficulty in breathing (DB), Nasal Discharge/Congestion (DC), Cough (C), Fever (F)

RSV Active surveillance

Notes to the FW for completion of the active surveillance form

Before you depart.
(i) Obtain list of children to be visited today + details of how to locate
(ii) Identify which of them have previously had a nasal washing in the last two weeks, and the lab result.
(iii) fill out date of active visits at top of the active surveillance form and date of next appointment (1 week or 1 month).
(iv) Check to take adequate materials/supplies for blood slides, nasal washings, and referrals.

At the house of the child
1. Remember to introduce yourself to the family/mother/head of household and identify yourself as from KEMRI for the RSV study. Remember to always be polite and to respect the wishes of the family.

2. Enter child's **KBC#** and **NAMEs/SEX** - take care to be accurate

MALARIA

FLWK - Ask mother if her child has had fever - hot body - in the last 7 days. This includes today.
TEMP - take axillary temperature using a Becton-Dickinson Thermometer (BD, New Jersey, USA) and record here.
Actions to take on the basis of these details are given in the manual.

NASAL WASHING

RSV2w - record here if child has had positive (see scenarios) RSV specimen in last 2 weeks. If so then we will not collect a further nasal washing today

BDWK / KEYWK - Now check mother for occurrence of difficulty in breathing and which of the key signs she has observed in her child over the last 7 days (including today). Note these are mother's observations, not the FW.

BDOBS / KEYOB - Observe child and record if s/he has difficulty in breathing now (allow child to settle before making observation). Observe child and record which of the key signs you can observe now (ie nasal discharge/congestion = *dc*; cough = *c*; hot body ie fever, = *f*). Note that these are FW observations, not what the mother indicates. Note that today, the mother could indicate the child has fever, but on inspection the FW cannot identify this. Hence record '*f* for question **KEYWK**, but do not include '*f* against KEYOB. One instance where this is different: the mother may indicate the child has a cough this day of observation, but you do not hear a cough. In this instance we assume the mother is correct since you may not be present long enough to hear a child cough.

NWREQ - A nasal washing is required if (at least) the following are indicated by either the mother OR the FW observation (i) the child has difficulty in breathing (*db*) OR (ii) the child has nasal discharge / congestion (*dc*), OR (iii) the child has cough (*c*). IMPORTANT - note

that these observations should be co-occurring to merit NW collection, i.e. it would not count if the mother observed a cough one day and hot body on another).

If NW required then carefully follow instructions in manual for collection.

REFERRAL

Brate - record the breathing rate as accurately as possible. Ensure child has settled before recording.

Assess if child has fast breathing (rate is 50 or more per minute, OR rate is 40 or more and child is older than one year ie 12 months or more in age).

Should remind mother of key signs and difficulty in breathing to decide when to bring child in again.

REFER - if child has fast breathing for age AND cough OR difficulty in breathing, then refer the child to the OP clinic for review (complete referral note in RSV exercise book), and provide mother with single journey bus fare to clinic. Emphasise that the mother should take the child today.

OTHER

AWAY - ask mother if the child has been out of the District since the last visit and, if yes, for how many FULL weeks.

Appendix G: Field Work Procedures

The following is the series of field study processes. It includes all the assessments of symptoms carried out in during active and passive surveillance and actions required (flow diagrams) details of which formed part of the Field worker manual not included in this thesis. Following these are the procedures for sample collection (nasal, oral-fluid and blood) used during the study.

Details of Flow Diagrams Included

RSV Active surveillance – weekly home visits by Field workers

Flow diagram 1- Assessment for fever and the requirement of blood slide (p 331)

Flow diagram 2- Assessment for respiratory signs and nasal wash requirement (p 332)

Flow diagram 3- Assessment for referral due to respiratory disease (p 333)

Follow-up visits

Flow diagram 4- Review for fever on last visit (p 334)

Flow diagram 5- Review for nasal wash result (p 335)

Out Patient Clinic Visits

Flow diagram 6 - Out patient clinic visit (p 336)

Standard Operating Procedures for Sample collection

Weekly visits by field worker to households

This is the main component of active surveillance of the study. On each visit the following assessments should be followed and prescribed actions taken (see appended flow diagrams). Appended also is an ACTIVE surveillance form that should be completed (notes accompany this form).

1. FEVER ASSESSMENT - IS A BLOOD SLIDE REQUIRED? See flow diagram **p 331**

- **If the child's measured temperature is raised >37.49°C:** Take **TWO** malaria slides and spot blood onto one filter paper. Complete **TWO** active surveillance forms. Give one slide and one form to the mother and take one back to the Unit. In addition, **refer mother to study clinic** (using exercise book) and **issue one-way bus fare.** Arrange to review child with result of film whether film +/-.
- **If history of fever during last 7 days but measured temperature not raised:** Take **ONE** malaria blood film and spot blood onto filter paper and take back to Unit for reading. Arrange to review child the following day only if film result is positive.

2. RESPIRATORY INFECTION - IS A NASAL SPECIMEN REQUIRED? – See flow diagram **G2 p 332**

- **if the child was POSITIVE for RSV within the last 2 weeks (14 days) -** no action is required, review next week.
- **if NEGATIVE (or EQUIVOCAL) for RSV over last 2 weeks -** the following questions must be asked:

  i) "Does your child have difficulty in breathing OR nasal-discharge/congestion OR cough today OR fever+ cough/nasal discharge?" Answer Y or N.

  ii) "Has he/she had difficulty in breathing OR any of the key symptoms since the last visit?" Answer Y or N

If the answer to question i) is **"Y"** a nasal specimen should be collected.

If the answer to question ii) is **"Y"** find out if the child has already had a nasal washing since the last visit. If not then a nasal sample should be collected.

Difficulty in breathing - signs

- nasal flaring
- use of accessory muscles, eg neck muscles
- indrawing (sucking in of lower bony chest wall on inspiration)

The key symptoms you need to be able to identify are:

Difficulty in breathing; cough (less than 30 days); fever (hot body or axillary >37.49C); nasal congestion/ discharge

If a nasal sample is collected then:

- Request that the mother monitors the child's health and if it does not improve or deteriorates (ie any of the symptoms worsen) then she should seek medical help from a local clinic or KDH.
- The nasal specimen should be transported back to the lab AS SOON AS POSSIBLE.

3. ACUTE RESPIRATORY DISEASE - IS REFERRAL REQUIRED? See Flow diagram **G3 p 333**

- **Does the child have COUGH (for less than 30 days) or BREATHING DIFFICLTY and does the child have FAST BREATHING?**

If the answer is "Y"

- The child should be referred to KDH research clinic for assessment. Issue referral note and one way bus fare to the clinic.
- A nasal washing should be taken, unless the child (i) has given a nasal wash nasal wash already this day OR (ii) was RSV positive within the last 2 weeks (14 days)

Consider a child to have fast breathing if:

breathing rate is ≥ 60 AND age < 2 months
breathing rate is ≥ 50 AND age < 12 months
OR
breathing rate is ≥ 40 AND age ≥ 12 months
(modification of WHO guidelines)

* You need to be able to accurately assess breathing rate.

Monthly visits of field worker to households

This activity is primarily to support the passive surveillance component of the study. During each monthly visit, the field worker will revisit the signs indicative of RSV infection with the mother as a general reminder of the study. The mothers will continuously be encouraged to present to KDH with an infant who has breathing difficulty or any two or more of the key symptoms (described earlier) consistent with RSV infection.

In addition, the three assessments made during weekly visits will also be carried out (as described above) ie, 1. Fever assessment - is a malaria slide required? 2. Respiratory infection - is a nasal specimen required? 3. Acute respiratory disease - is referral required. (note: read 'last week' for 'last visit').

Oral-fluid will be collected every three months throughout the study period. In some instances venous blood will also be collected at the same time.

Passive visits or Active referral to the OP Clinic

Infants may be brought to the OP Clinic by parents either as part of the passive referral procedure or after referral from the field for one reason or another. Having ascertained the reason why the child is presenting, the relevant clinic form should be completed, and child reviewed by clinical officer (CO). See appended 'RSV study sick child clinic visit' form.

Child follow up

1. CHILD WITH FEVER - see flow diagram **G4 p 334**

ARRANGEMENTS FOR FOLLOW UP AND TREATMENT

- **All malaria slides will be examined at the unit on the same day.**
- **All children with positive slides OR children with current fever (>37.49C) should be reviewed the next day.** If a child seen at routine follow-up has a measured fever >37.49°C s/he will be expected at the clinic the following day. If they attend clinic they will receive definitive assessment and treatment. Active surveillance hospital form will be completed in clinic. If they fail to attend the clinic as arranged, they should be reviewed in the field the next working day.

Review in the field

Slide positive

- Measure temperature and completed follow-up section of form.
- **Administer malaria drugs.**
- Advise parents that they **must take the child to a clinic (preferably KEMRI) or other outlet if s/he deteriorates or fails to improve within 1-2 days.**
- Make it clear that if the family chooses to bring the child to KEMRI clinic, the fare will be refunded.
- Visit routinely the following week/month according to time of year.

Slide negative

- **If child no longer gives history of fever and is improving** – no further action.
- **If child remains unwell or febrile (either measured or on history)** – repeat slide and filter paper and refer to KEMRI OPD (issue one-way bus fare).

2. CHILD GIVING NASAL SPECIMEN - see flow diagram **G5 p 335**

- If a nasal washing was RSV **negative** then the result is reported to the mother at the next visit.
- If the nasal washing result is RSV **positive** or **equivocal,** the field worker will be required to revisit the child AS SOON AS POSSIBLE to collect (i) acute blood (2ml) (ii) oral-fluid and (iii) make an appointment to return after one month to collect convalescent blood and oral fluid.

3. CHILD REFERRED WITH ACUTE RESPIRATORY DISEASE - see flow diagram **G6 p 336**

At the OPD research clinic, a clinical officer will review the child.

- If symptoms of ARI are serious involving risk or danger signs indicative of severe or v. severe pneumonia, the child will be admitted to Ward 1.
- If symptoms of ARI are moderate or mild, then home (symptomatic) treatment will be recommended. Request that the mother monitors the child's health and if it does not improve or deteriorates (any of the symptoms worsens) then she should seek medical help from a local clinic or KDH. A review will be made by the field worker at the following visit.

A note on making appointments

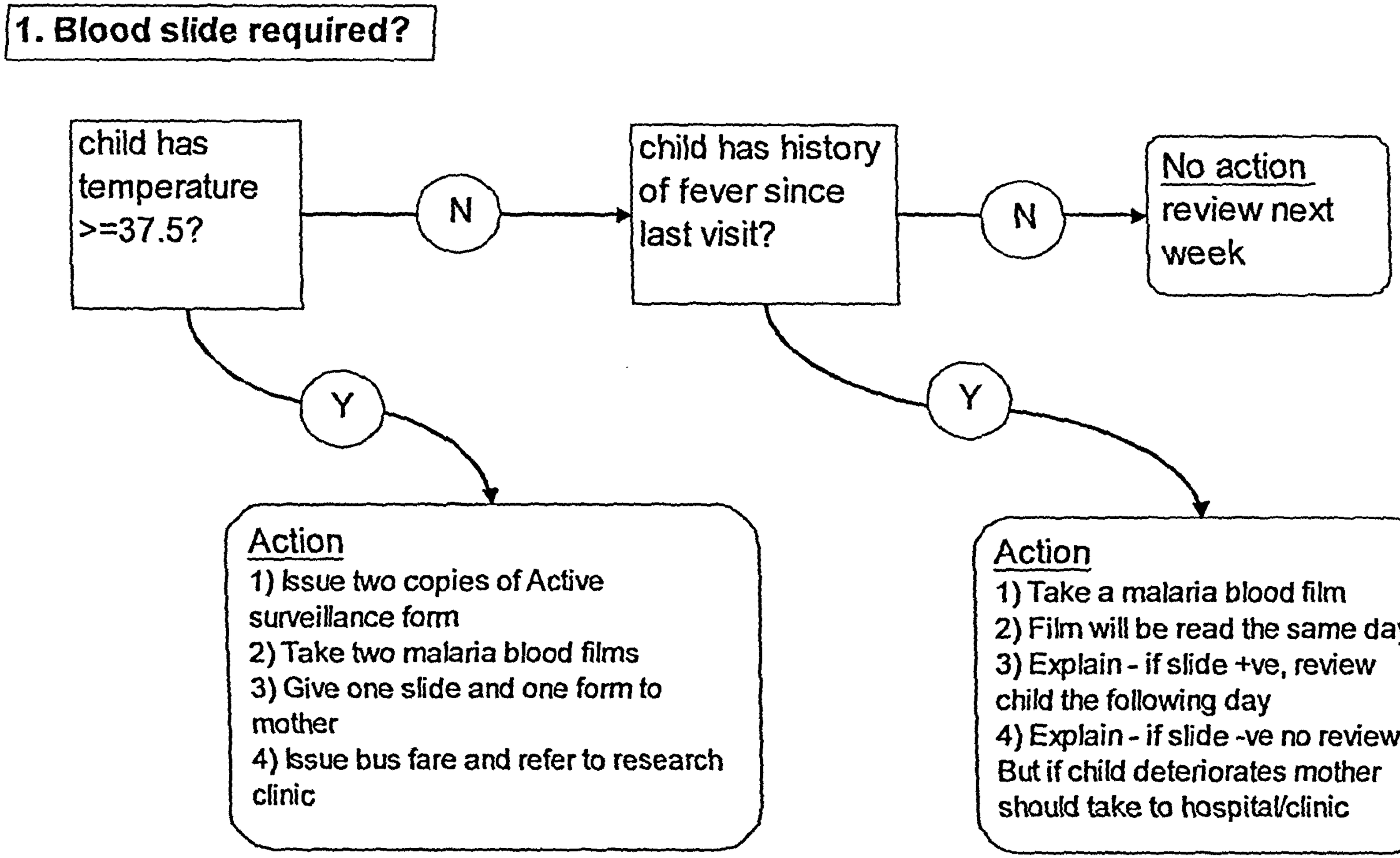
Mothers will be bringing their infants/children in to KDH for routine 3 monthly blood sampling as part of the main cohort and for vaccination appointments at the MCH clinic. Wherever possible make RSV appointments at the research clinic rather than at the home to coincide with these visits.

Accuracy in data collection and storage of samples.

Accuracy in data collection is a top priority. The following are some simple guidelines:

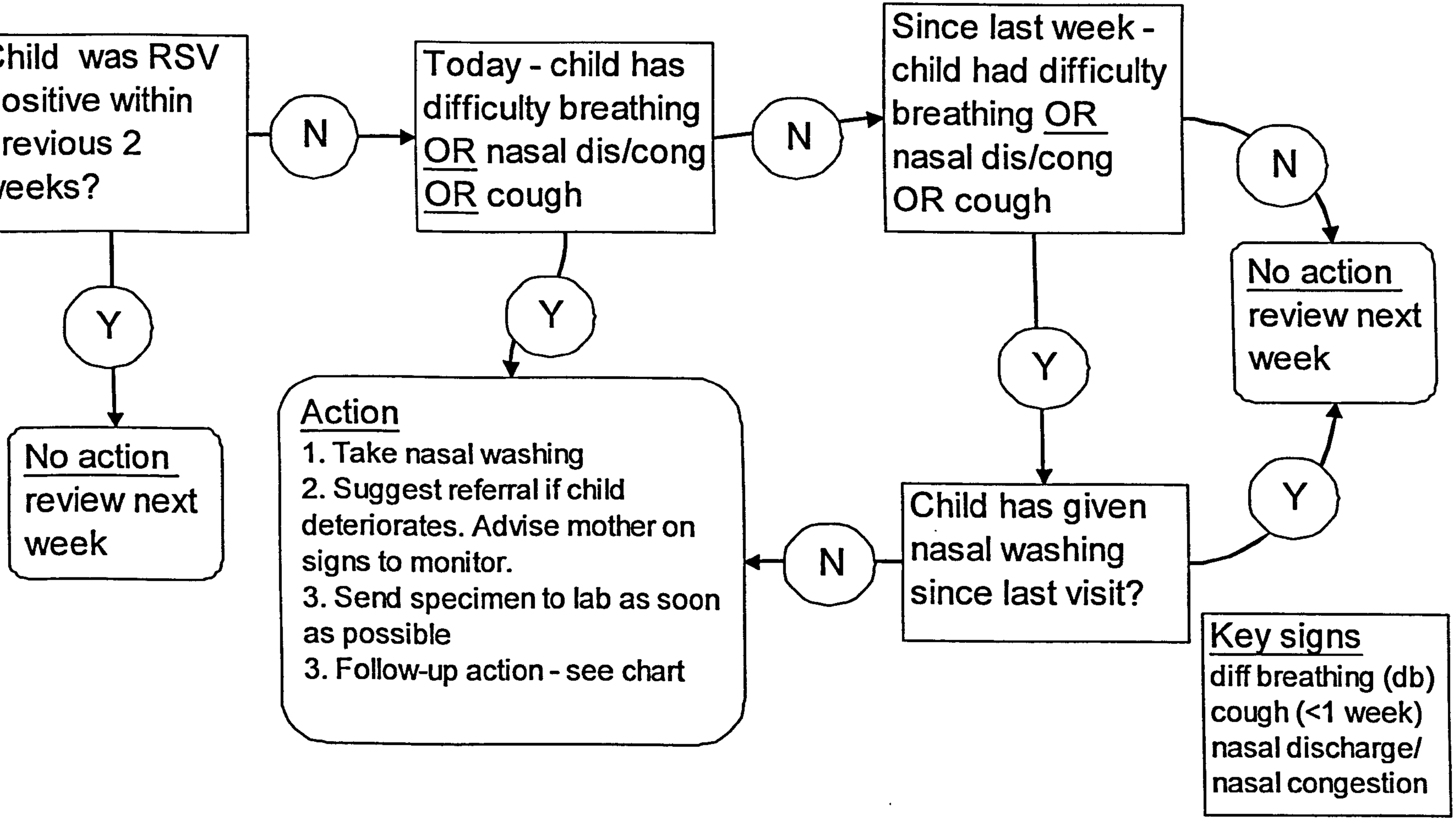
- Label all bottles and tubes for saliva, nasal and blood collection BEFORE TAKING A SAMPLE – it is easy to forget to label materials in the noise and confusion of sample taking. All samples should be labelled with the individual's name, serial number and the date of collection of the sample.
- Make every effort to return collected specimens to the research unit as quickly as possible, maintaining cold conditions as required.
- Complete forms and questionnaires carefully. Complete all sections required. Do not be hurried. Do not make assumptions about answers elicited - if unsure, return to the question and ask again.
- Be methodical in your actions. Make sure you are clear exactly what you intend to do before arriving at the household.
- On return to the unit (i) go through the forms to ensure that there are no blatant mistakes and hand in the forms to the field co-ordinator. (ii) Relate all the problems encountered in the field to the field co-ordinator and also note them down for discussion. (iii) Enter data onto computer, or provide to data entry clerk, as soon as you can, (iv) enter kbc# into day book in OP clinic of individuals from whom nasal washing s have been collected (update this as and when the NW result is known, and make appointments as necessary)/

Appendix G1: Flow Diagram 1

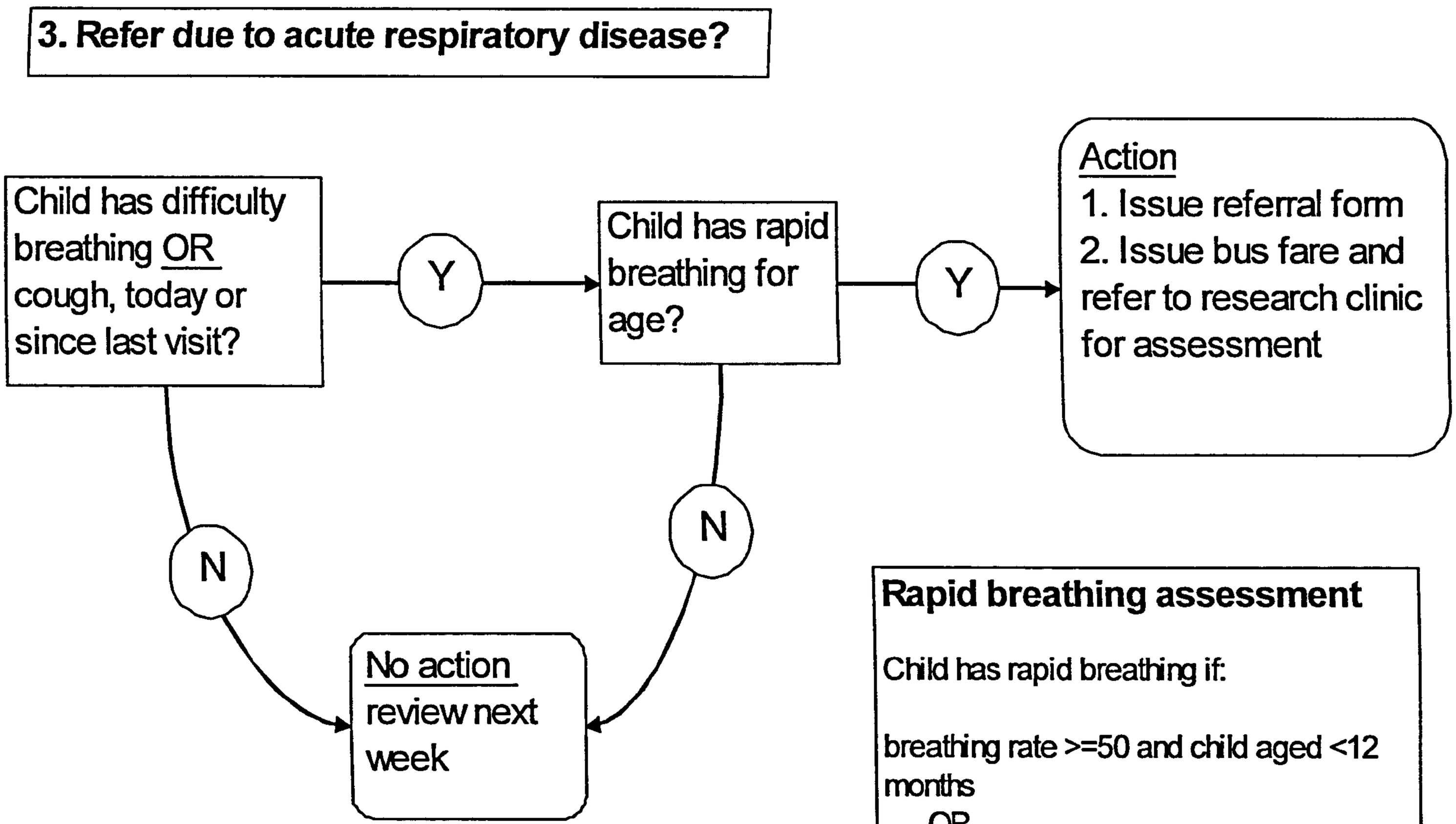


Appendix G2: Flow Diagram 2

2. Nasal washing required?



Appendix G3: Flow Diagram 3



Rapid breathing assessment

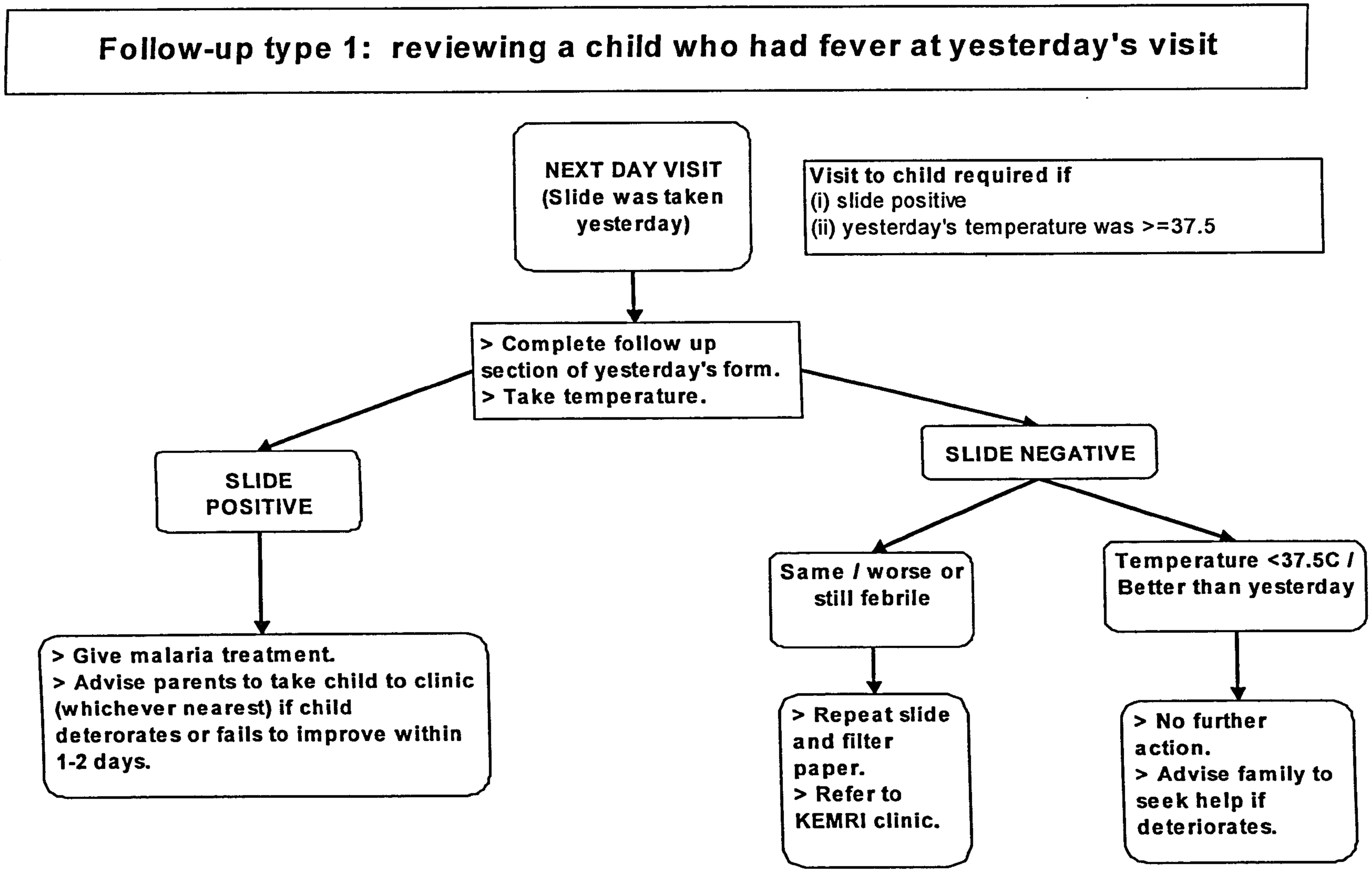
Child has rapid breathing if:

breathing rate >=50 and child aged <12 months

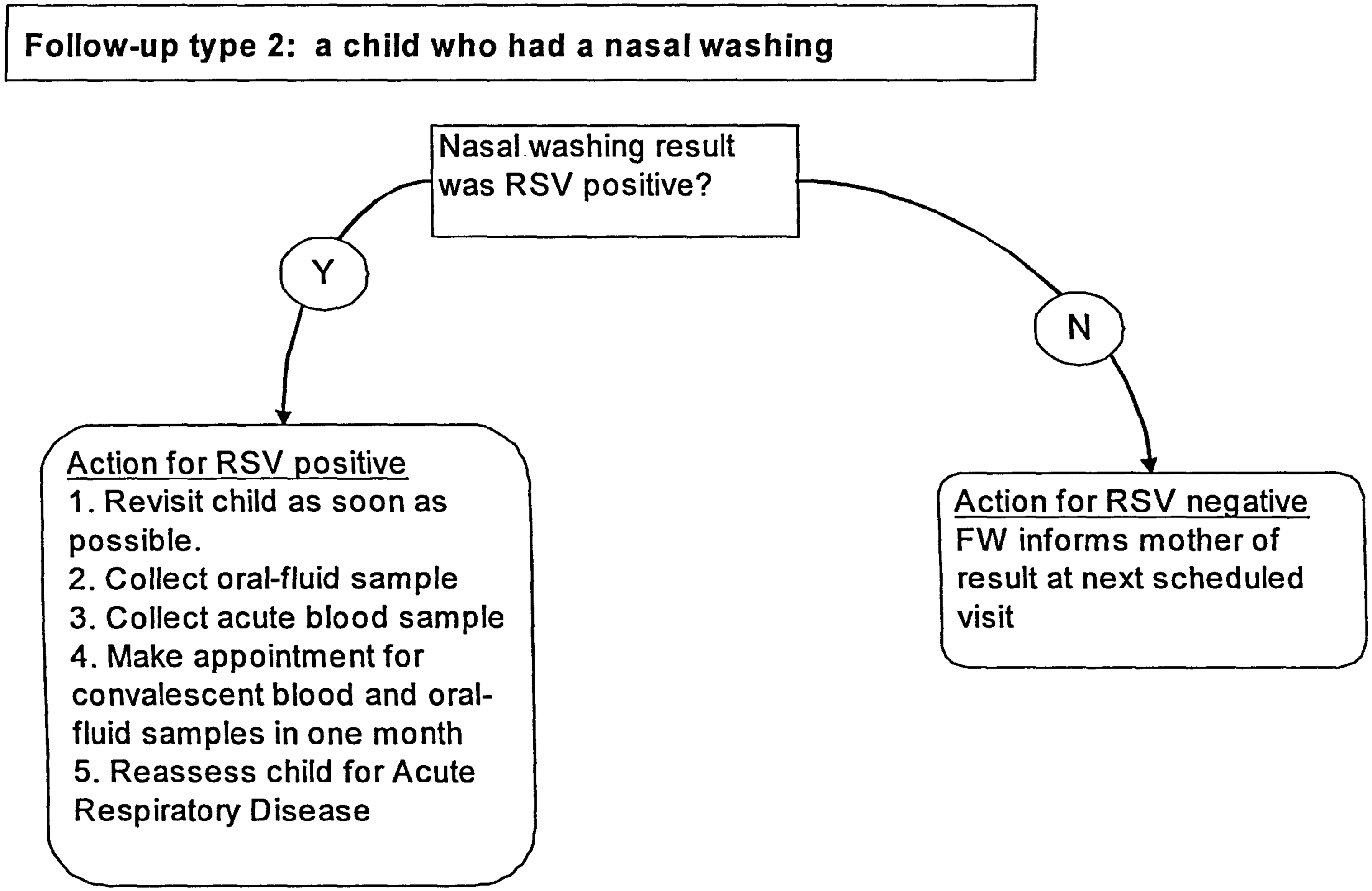
OR

breathing rate >=40 and achild aged >=12 months

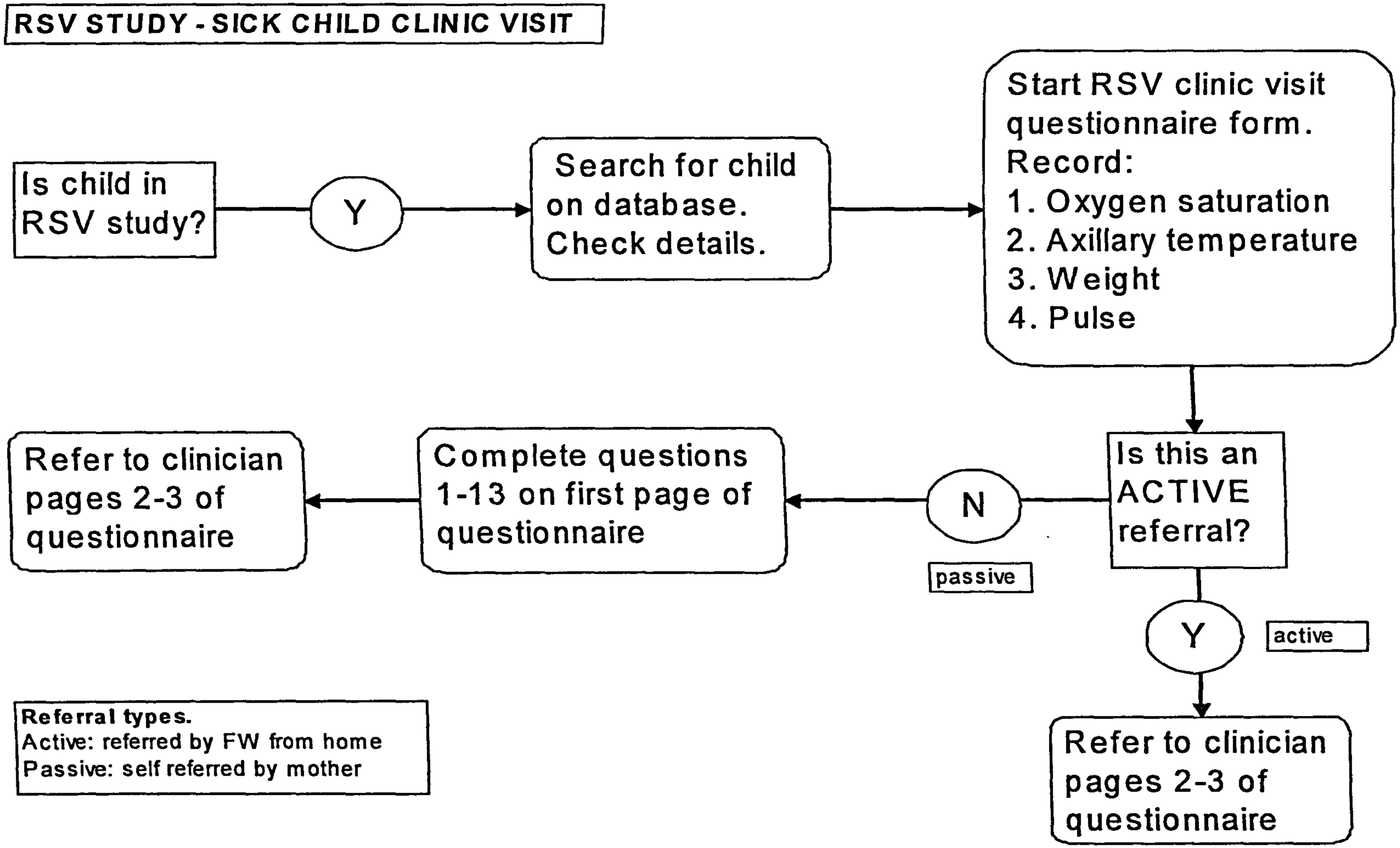
Appendix G4: Flow Diagram 4



Appendix G5: Flow Diagram 5



Appendix G6: Flow Diagram 6



Sample Collection Procedures

Temperature recording

All temperatures will be taken from the armpit (axillary temperatures) using a Becton-Dickinson Digital Thermometer (New Jersey, USA). Do not round off the figures on the thermometer but record it as it is on the forms.

Recording breathing rates

Defined as the number of breaths per minute. Obtained by observing the frequency of inspiratory phase over 30 seconds then multiply by 2.

Collecting nasal specimens

(i) Nasal washing using blue bulb - observe short video presentation

1. Requirements

- RSV sample/screen form
- Normal saline
- 1oz (30ml) Davol bulb syringe (Abbort, Ashland, Ohio-USA)
- Collection bijou/universal/medicine cup
- Marker pen
- Disposable gloves/paper towel.

2. Procedure

2.1 Wash hands (if possible) and replace gloves

2.2 Explain procedure to the mother/carer

2.3 Pour required quantity of saline into collection container -

child <6m 5ml

child 6-11m 7ml

child 12+m 10ml

child >5yr 15ml

adult 20ml

label with date, name and KBC or RSV number.

2.4. Have the mother hold child from behind in upright position and stabilize head (neutral or slightly extended)

2.5 Suction saline into bulb. Insert bulb into one nostril (holding bulb horizontally) until nares occluded (you should be firm and deform the nose slightly). Hold collection container under other nostril Tip bulb slightly towards patient, instill firmly all fluid into nose using

thumb action. Fluid should eject from other nostril. Suction residual mucus from both nostrils into the bulb syringe and mix with the saline. Transfer residual fluid from bulb into collecting container.

2.6 If no or little fluid emits from other nostril, or if less than 1 ml of fluid recovered, replenish fluid to correct amount (see above) in container, suck into bulb, and repeat procedure in alternative nostril.

IMPORTANT: It is crucial that you make every effort to have fluid pass out of the alternative nostril. In this way we know fluid has gone to the nasopharynx where the virus replicates.

2.7 Dispose of gloves and wash hands.

3. Storage of specimens

3.1 Immediately after collection, and securing screw cap lid tightly, put the specimen in the icebox. Transport to lab as soon as possible

3.2 Complete Microbiology request form.

IMPORTANT: Contamination between containers, syringes used to withdraw saline, gloves, and bulbs, should be avoided. We will be using very sensitive techniques to detect the virus and any contamination will be detectable. So, DO dispose of gloves, and store separately used bulbs, after each procedure, make sure lids of vials are in firmly, DO NOT reuse bijous/universals.

Oral fluid collection

1. Requirements

- RSV sample/screen form
- Indelible marker pen
- Disposable gloves and paper towel
- ORACOL collection device (Malvern Medical Developments, Worcester, UK)

2. Procedure

2.1 Wash hands and replace gloves

2.2 Explain procedure to the mother/carer

2.3 Remove ORACOL device from covering and transport tube

2.4 Holding by stick rub sponge gently but firmly around gums (front and back, upper and lower, including base) for 1 minute.

IMPORTANT - the sponge must be fully saturated with oral fluid and so one minute of use is required.

2.5 Replace device into stoppered tube (sponge at bottom), label with KBC number, date, 'RSV-NO'.

2.6 Dispose of gloves and wash hands

3. Storage of specimens

3.1 Immediately after collection, put the specimen in the icebox. Transport to lab as soon as possible

3.2 Complete Microbiology request form.

Collecting venous blood

1.Venepuncture

1.1 Explain procedure to parent/guardian.

1.2 Ask the mother/guardian to help you hold the child.

1.3 Select an appropriate site and apply a tourniquet beyond.

1.4 Clean the site with spirit and leave to dry.

1.5 Identify vein visually or by palpation and introduce the needle or butterfly.

1.6 Withdraw blood.

1.7 Apply swab, press firmly and withdraw needle/butterfly.

1.8 Check that there is no bleeding from the VP site.

Problems

Inability to draw blood?

- Needle not in the vein (commonest).
- Needle in vein but not properly positioned.
- A clot may be blocking the needle.
- Syringe may be leaking.

Risks of taking blood

To Fieldworker

- Accidental needle stick injuries may transmit diseases like HIV, Hepatitis.

To child

- Pain at the VP site

- Bleeding and swelling of the VP site.
- Infection at VP site which may spread elsewhere.

Appendix H: Sick child clinic visit questionnaire

RSV STUDY - SICK CHILD CLINIC VISIT QUESTIONNAIRE

FIELD WOKER

complete page 1 + record Oxygen Sat/Pulse/Weight/Axillary temp on page 2

KBC # [ ][ ][ ][ ][ ]

RSV # [ ][ ][ ]

Child Name [ ] [ ] [ ]

Date of Birth [ ][ ]/[ ][ ]/[ ] [ ] dd/mm/yy

Date today [ ][ ]/[ ][ ]/[ ] [ ] dd/mm/yy

Visit type [ ] active (referral by FW) [ ] passive (self-referral) - Tick one box

---

***IF ACTIVE REFERRAL** (ie referred by FW from home) - Refer to clinician to complete only pages 2-3*

IF PASSIVE REFERRAL (ie mother self referred) - FW complete the below

1. Ask mother - has child had a fever in the last 7 days? [ ] y/n
2. Has the child a temperature >37.49C (see page 2) [ ] y/n

If either 1 or 2 is 'yes' then do blood slide.

3. Blood slide taken [ ] y/n
4. Blood slide result [ ] +/-
5. Ask mother - has the child had difficulty in breathing in last 7 days? [ ] y/n
6. Ask mother - which of the key symptoms in the last 7 days? [ ]
7. Observe the child to have difficulty breathing today? [ ] y/n
8. Observe which key symptoms today ? [ ]
9. Is a nasal washing required? [ ] y/n
10. NW time [ ][ ]:[ ][ ] hh:mm
11. NW result [ ] +/-
12. Breathing rate (cycles / min) [ ][ ]
13. Does child have fast breathing for age [ ] y/n
14. Refund single (whole shillings) [ ][ ] ksh

Key symptoms: cough c, fever f, nasal discharge/congestion dc

Now refer child to clinician .....next page

To be completed by CLINICIAN

HISTORY

How long has the child been sick? .....[ ][ ] DAYS

What is/are the main problems?(Please answer Y/N)

FEVER.....[ ]

COUGH.....[ ]

RUNNING NOSE/NASAL DISCHARGE/NASAL CONGESTION.....[ ]

DIFFICULTY IN BREATHING.....[ ]

OTHERS(specify below).....[ ]

PHYSICAL EXAMINATION

OXYGEN SATURATION	[ ][ ][ ]%
AXILLARY TEMPERATURE	[ ][ ].[ ][ ]C
WEIGHT	[ ][ ].[ ][ ]kg
HEART RATE	[ ][ ][ ]/min

FW to complete

COUGH HEARD	[ ] Y/N
HOARSENESS	[ ] Y/N
SORETHROAT	[ ] Y/N
RHINITIS	[ ] Y/N
PHARYNGITIS	[ ] Y/N
CYANOSIS	[ ] Y/N
EAR INFECTION	[ ] Y/N
CONJUCTIVITIS	[ ] Y/N
PROSTRATED	[ ] Y/N
RESPIRATORY RATE	[ ][ ]/min
NASAL FLARING	[ ] Y/N
INDRAWING	[ ] Y/N
CRACKLES	[ ] Y/N
WHEEZES	[ ] Y/N

STRIDOR [ ] Y/N

DIAGNOSIS

#Indicate if any of these:
LRTI / URTI
Bronchiolitis
pneumonia - mild / severe / v. severe

PLAN

*Nasal washing required y/n [ ]
Admit child y/n [ ]

TREATMENT

*** Note: Request nasal washing (NW) if (i) child has nasal discharge/congestion OR (ii) cough + fever OR (iii) difficulty in breathing OR (iv) has any of respiratory disease diagnoses given above# AND FW has not already taken NW.**

Notes to the FW for completion of the RSV cohort child visit proforma

The FW is to enter information on pages 1.
Also the FW will record (i) Oxygen saturation (ii) Axillary temperature, (iii) Weight, (iv) Pulse, and enter these measurements on page 2 of the form.

Page 1
COHORT NUMBERS
Record KBC and RSV numbers - take care to ensure accuracy.

CHILD DETAILS
Record other child details - names and date of birth - accurately.

DATE TODAY
Record date of interview/visit.

REFERRAL type
Record 'ACTIVE' if this is a referral by FW from child's home.
Record 'PASSIVE' if mother/child self-referred to clinic.
If ACTIVE, then (having taken measurements for page 2) pass child onto clinician for medical review (page 1 will have been collected at the child's home).
If PASSIVE, then complete remaining part of page 1.

PASSIVE visit
1. Ask mother if her child has had fever - hot body - in the last 7 days. This includes today.

2. Answer 'y' if the temperature is 37.5C or more.
If answer to 1. or 2. is 'y' then take a malaria blood slide.

3-4. Answer these questions relating to collection of blood slide and recording result.

5-6. Now check mother for occurrence of difficulty in breathing and which of the key signs she has observed in her child over the last 7 days (including today). Note these are mother's observations, not the FW. Note that today, the mother could indicate the child has fever, but on inspection the FW cannot identify this. Hence record 'f' for question 6, but do not include in the answer to question 8.

7-8. Observe child and record if s/he has difficulty in breathing now (allow child to settle before making observation). Observe child and record which of the key signs you can observe now (ie nasal discharge/congestion = dc; cough = c; hot body ie fever, = f. Note that 7-8 are FW observations, not what the mother indicates.

9. A nasal washing is required if (at least) the following are indicated by either the mother or the FW observation (i) the child has difficulty in breathing (db) OR (ii) the child has nasal discharge/congestion, OR (iii) the child has fever + cough.

10-11. Complete if nasal washing is required. Make sure the sample is properly labelled and sent to Micro lab with request form.

12. Record breathing rate as accurately as possible. Ensure child has settled before recording.

13. Assess if child has fast breathing. 'y' if rate is 50 or more per minute. 'y' if rate is 40 or more and child is one year or older ie 12 months or more in age.

Should remind mother of key signs and difficulty in breathing to decide when to bring child in again.

Appendix I: Limitation on provision of services

RSV COHORT STUDY

LIMITATIONS ON PROVISION OF SERVICES

CHILDREN PRESENTING TO OPD CLINIC - POSSIBLE SCENARIOS:

1. Child presents with symptoms/signs consistent with an acute respiratory infection (upper or lower) as defined by study*.

   ACTION: (i) review by CO, (ii) free treatment#, (iii) refund travel cost@.

2. Child presents with fever only (which is consistent with malaria)

   ACTION: (i) review by CO, (ii) blood film as appropriate, (iii) free treatment, (iv) refund travel cost.

3. Child presents with other illness e.g. eye/ear infection, wounds, scabies, jaundice, oral thrush etc.

   ACTION: (i) review by CO, (ii) free treatment (iii) NO travel fare refund.

4. RSV child is accompanied by a sick sibling, or mother is sick and not in HH study.

   ACTION: (i) review by CO (ii) prescription/treatment paid for by the mother, (iii) NO travel fare refund (not applicable if RSV child is sick also).

5. RSV child, on being seen by the clinician, is referred to the ward.

   ACTION: the study will not cover any additional costs that may result.

Definitions:

**History of ARI symptoms* - Difficulty in breathing, nasal discharge/congestion OR cough + fever. However, in practice, any sign of respiratory infection is sufficient to merit free treatment and travel costs.

#*Free treatment* - free provision of prescribed drugs routinely dispensed from OPD. If the required drug is not available at OPD, the mother is advised that she will have to pay for the drug, or she can take an alternative (less appropriate) drug from OPD free of charge.

@*Travel cost* - reasonable return travel fare.

Any difficulties contact: Emelda Okiro (x126/107) or James Nokes (x131)

1. Laboratory Procedures Appendices

The attached appendices give details of Laboratory procedures used during the study. The first set of appendices (Appendix J and K) details the procedures for handling nasal specimen. Following this is the SOP used to perform the RT-PCR (Appendix L). The next few appendices provide specific details on standard operating procedures for both the IgG capture ELISA and the RSV-specific ELISA used to process the oral-fluid samples.

Appendix J: Cytospin method for cell separation

SOP: CYTOSPIN METHOD FOR CELL SEPARATION

Date of implementation_____ Date of review_____

Signed by_____ Signed by_____

The cytospin method is another technique of depositing cells onto slides for IFA. This rapid method is used to detect viral antigens in cellular specimens such as nasopharyngeal aspirates and for the identification and typing of virus isolates in tissue culture cells.
Nasopharyngeal aspirates should be processed first to deposit and concentrate the cells according to SOP: Manual Cell Separation.

Procedure

- Prepare the cytofunnel assembly. Ensure the slide is right way around and filter is positioned correctly (for re-usable funnels).
- Load the cytofunnel assembly into the cytospin centrifuge bowl.
- Inoculate 200µl of the cell suspension (Nasal wash bulb or Nasal wash syringe) into the cytofunnel and close the funnel with the lid provided, in the safety cabinet.

  NOTE: Do not allow funnel to tilt allowing fluid to contact filter or to touch sides of the funnel during inoculation.
- For NPA first perform one cell suspension wash using PBS (refer to SOP: Manual Cell Separation) then proceed to inoculate 150µl of the cell suspension into the cytofunnel.
- Balance the bowl and replace lid before removing from the cabinet.
- Load centrifuge bowl into centrifuge.
- Spin at 1000rpm for 10 minutes.
- After spinning transfer to safety cabinet and remove the slide from the cytofunnel assembly and discard the funnel (if disposable) or place the funnel in Virkon solution (if re-usable funnel).
- Air dry the slide and fix the cells in 100% acetone @ 4°C for 15 minutes.
- The cells are then ready for IFAT staining.

Appendix K: The detection of Respiratory Syncytial Virus (RSV) using the direct immunofluorescence test.

SOP: THE DETECTION OF RESPIRATORY SYNCYTIAL VIRUS (RSV) USING THE DIRECT IMMUNOFLUORESCENCE TEST.

Date of implementation_____ Date of review_____

Signed by_____ Signed by_____

PURPOSE

Light Diagnostics Respiratory Syncytial Virus DFA Kit is intended for the detection and identification of RSV in inoculated cell cultures and direct respiratory specimen cell preparations.

PRINCIPLE OF THE TEST

Light Diagnostics RSV DFA kit utilizes a direct immunofluorescent antibody technique for identifying RSV in cell preparations made from nasopharyngeal aspirates (NPA). The monoclonal antibodies are labelled with FITC (fluorescein isothiocyanate), which fluoresce apple - green when illuminated with ultraviolet light. The labelled antibody will bind to viral antigen present in the specimen. Unbound reagent is removed by washing with buffer. Cells in positive specimens will fluoresce apple – green while uninfected cells will stain dull red due to the presence of Evans blue.

MATERIALS

- The recommended respiratory sample is nasopharyngeal aspirate which, when correctly collected, should provide large numbers of respiratory epithelial cells. Nasopharyngeal aspirate/nasal washing will be obtained as per protocol KEMRI/WTRL/
- Acetone – to be stored at 4^0C
- Phosphate buffered saline (PBS) pH 7.5 (prepared as per instructions in the kit).
- Teflon coated microscope slides with single 6mm well (Shandon cytoslides)
- Control slides
- Precision pipette to deliver 25µl.
- Wash baths
- Wide bore disposable pipettes
- Decontaminant – Virkon
- Cover slips.
- Moist incubator at 37^0C.
- Low speed centrifuge.

- Epifluorescence microscope with filter system for FITC (maximum excitation wavelength 490nm, mean emission wavelength 520nm) and x200-x400 magnification.

The sample will be brought from the ward / OPD to the microbiology department. Upon receipt in microbiology laboratory the following steps will be performed:

CELL SEPARATION

- Place the sample at 4^0C until washing is initiated. The specimen should, however, be processed immediately.
- All manipulations of specimen should be done in a safety cabinet and discarding of material should be done in a disinfectant (Virkon).
- Add 6ml PBS to the sample prior to centrifugation in disposable centrifuge tube (blue top) to reduce the viscosity and break the mucus.
- Centrifuge at room temperature (15^0- 30^0C) for 10minutes at 1800rpm.
- Discard supernatant. Suspend the cell deposit in 6ml PBS and gently pipette the cells up and down with a wide bore pipette, until the mucus is broken up and cellular material released. Avoid vigorous pipetting / vortexing to prevent damage to the cells. .
- Remove and discard any visible flecks of mucus remaining at this point. Excess mucus must be removed as it will prevent adequate penetration of the reagent and may result in non-specific fluorescence.
- Centrifuge at room temperature (15^0- 30^0C) for 10minutes at 1800rpm.
- Discard the supernatant and resuspend in a little PBS depending on the visual assessment of the suspension (not too dilute and not too cloudy).
- Add 25µl of the suspension on the cytoslide.
- Air-dry the slide.
- Fix the slide in cold acetone at 4^0C for 10 minutes.
- Dry the slides completely at room temperature and store at -70^0C until ready for immunofluorescent assay.
- Pipette the remaining cell suspension into a vial and add an equal amount of freezing mixture. Label **RSV NWB-2** and store at -70^0C.

PROCEDURE

- Place a drop of *Light Diagnostics RSV DFA* reagent to the fixed cell preparation and to a positive control slide. Ensure that the reagent covers the entire well area.
- Incubate the slides with reagent for 30 minutes at 37^{0}C in a moist chamber. **Do not** allow the reagent to dry on the specimen, as this will cause the appearance of non-specific staining.
- Wash off excess reagent in a fresh change of PBS.
- Drain off PBS and allow the slide to air dry at room temperature.
- Add one drop of mounting fluid to the centre of each well and place a coverslip over the mounting fluid and specimen ensuring that no air bubbles are trapped.
- Examine the entire 6mm well area containing the stained specimen using an epifluorescence microscope. Fluorescence should be visible at X200-X500 magnification. (For best results specimens should be read immediately after staining, but specimens may be stored at 2- 8^{0}C in the dark for up to 72hours).

INTERPRETATION OF RESULTS

1. Control slides

RSV-infected cells will show apple-green fluorescence in the cytoplasm and associated with syncytia.
The negative control well should show cells staining a dull red.

2. Clinical specimens

- Apple-green fluorescent intracellular cytoplasmic granules are seen in respiratory epithelial cells infected with RSV.
- In later stages of infection, RSV antigen may be restricted to isolated areas in the cytoplasm appearing as small ill-defined granules singly or in clusters.
- Uninfected cells stain with the Evans blue counter stain and appear red.

Interpretation

- **Positive diagnosis** is made when one or more cells show typical fluorescence in the fixed, stained specimen.
- **Negative diagnosis** is made when fixed, stained specimens do not exhibit fluorescence after staining with the reagent.
- For directly stained nasopharyngeal aspirate specimens at least 20 uninfected respiratory epithelial cells must be visible within the slide well before a negative result is reported.
- A sample containing fewer than 20 cells is considered inadequate, and the test invalid.

QUALITY CONTROL

- **Viewing slides by second person blind to first**
  A second person should examine the slides without knowing the results of the first person then compare the two results. Results found to differ between the technicians should be reassessed and consensus obtained.

- **Checking procedure and quality of slides - cell count, staining, etc**
  Always check that there are at least 20 cells in the entire preparation on the slide. The stain should not dry on the slide and so always make sure there is water in the incubator (i.e. there's enough moisture in the incubator). Proper washing of the slides after staining should be ensured – make sure there is fresh PBS every time slides are to be washed.
- **Equivocal result procedure**
  If two people get equivocal results on the same slide, then it should be reported so. If one says equivocal and the other a different result, then the two should come to an agreement on what result to give to the clinician.
- **Set of slides with cell dilutions**
  These should be used once every month to Q.C the entire staining procedure as an internal quality assessment scheme.
- **Sending slides to the UK for second opinion**
  This should be at least once in every 3months. There should be at least 10 of each (i.e. 10 positives, 10 negatives and 10 equivocal).

REAGENT PREPARATION

1. **Freezing mixture for the NPA**

   - 20% Foetal Calf Serum (FCS)
   - RPMI
   - Antibiotics

2. **Freezing mixture for saliva**

   - PBS
   - 20%Foetal calf Serum (FCS)
   - 0.02% Sodium Azide

3. **Phosphate Buffered Saline (PBS)**

   - Packet of PBS salts in 1 litre distilled / deionised water.
   - 10ml Tween 20 / Sodium Azide Solution

Appendix L: RSV protocols for RT-PCR

RSV protocols for RT-PCR from random primer cDNA for multiplex and nasal washing samples (due to low levels of RNA)

1) Extraction of RNA from samples
   - This protocol uses the Qiagen Viral RNA extraction kit and follows the protocol in the manufacturers handbook (supplied with the kit) with the exception that the addition of the carrier RNA to buffer AVL is not done. Also the optional spin in the handbook (step 9) is not necessary.

2) To make cDNA for use in RSV multiplex and RT-PCR from nasal washings (make up reagent mix with exception of RNA in the clean room, do not take RNA into this room).
   - Using the Qiagen Omniscript Reverse Transcription kit, all of the reagents (10X buffer, 5mM dNTPs, 250ng/ul random primers (diluted from 1ug/ul stock 1 in 4) and nuclease free water) are thawed and vortexed to mix prior to use.
   - Label 0.5ml tubes with the sample names and also label the tubes with "rp cDNA" so that you know what the samples are if you need to go back to them at a later date.
   - The reagents are added to one 0.5ml tube according to the following table, resulting in a mix for the number of samples that will be used (e.g. if working with 5 samples and a negative control, then the volumes will be multiplied by 6):

Reagent	Volume per sample	Final concentration
10X RT Buffer	4.0ul	1X
5mM dNTPs*	4.0ul	
250ng/ul random primers	1.0ul	
Omniscript RT enzyme	1.0ul	
Water (provided)	10.0ul	

*Note: dNTPs provided with the kit will run out quickly using this method, therefore separate dNTPs will need to be purchased and made up at the same concentration for use to prevent wastage of the kit.

   - The contents of the tube are mixed by vortexing and then 20ul aliquoted into each of the sample tubes.
   - The caps on the tubes are then closed and taken to the main lab, where 20ul of RNA for each sample is added to the corresponding tube of cDNA reagents.
   - The tubes are then capped and incubated at 37°C for 1 hour and then either used for PCR directly, or frozen until needed at -20°C.

3) RSV/Flu multiplex RT-PCR from random primer cDNA

A) Primary PCR

- Start PCR block on program 19 (94°C for 2 minutes; 35 cycles of - 94°C for 1 minute, 50°C for 1 minute, 72°C for 1 minute; extension at 72°C for 5 minutes and cooled)
- In clean room, take out Qiagen *Taq* mastermix, primary multiplex primer mix (5uM)(see attached sheet for sequences), 25mM $MgCl_2$ and water vials and defrost.
- Once defrosted, add reagents as follows per sample (except for cDNA – done in main lab prior to running PCR):

Reagent	**Volume per sample**
PCR Mastermix	50ul
25mM $MgCl_2$	8ul
5uM first round primers	1ul
Water	21ul
CDNA to be added in main lab	*20ul*

- Aliquot 80ul of this PCR mix to separate labelled tubes for the samples
- Take to main lab and add 20ul cDNA to each tube from the corresponding sample.
- Vortex to mix and put on PCR machine and start the programme (enter key on machine) and leave to allow PCR reaction to complete (about 4 hours).
- Once complete, proceed directly with secondary PCR and then freeze PCR samples for future use in the N gene PCR.

Secondary PCR

- Run program 20 of PCR block (94°C for 2 minutes; 35 cycles of - 94°C for 1 minute, 60°C for 1 minute, 72°C for 1 minute; extension at 72°C for 5 minutes and cooled). Also defrost all reagents (mastermix, water, $MgCl_2$ and second round primer mix (2.5uM))
- Once defrosted, add reagents as follows per sample (except for first round product – done in main lab prior to running PCR):

Reagent	**Volume per sample**
PCR Mastermix	25ul
25mM $MgCl_2$	4ul
5uM first round primers	10ul
Water	11ul
1st round PCR product to be added in main lab	*2ul*

- Aliquot 48ul of this PCR mix to separate labelled tubes for the samples
- Take to main lab and add 2ul first round PCR product to each tube from the corresponding sample.

- Vortex to mix and put on PCR machine and start the programme (enter key on machine) and leave to allow PCR reaction to complete (about 4 hours).
- Once complete, proceed directly with detection of products by gel electrophoresis or freeze PCR samples for analysis next day.

4) N Gene RT-PCR

(Note: this is carried out as a secondary PCR on the first round products from the RSV/Flu multiplex protocol)

- Start PCR block on program 11 (95°C for 2 minutes; 30 cycles of - 95°C for 45 seconds, 54°C for 45 seconds, 72°C for 1 minute; extension at 72°C for 5 minutes and cooled)
- In clean room, take out Qiagen *Taq* mastermix, N1+N2 primer mix (10pmol/ul – made with 40ul water and 5ul each of N1 and N2 100pmol/ul stocks) and water vials and defrost.
- Add 25ul mastermix, 23ul water and 1ul N1+N2 primer mix per sample to a tube and mix.
- Aliquot 49ul of this PCR mix to separate labelled tubes for the samples
- Take to main lab and add 1ul of the first round PCR products from the multiplex analysis above to each tube from the corresponding sample.
- Vortex to mix and put on PCR machine and start the programme (enter key on machine) and leave to allow PCR reaction to complete (about 3-4 hours).
- Once complete, proceed directly with detection procedure and then freeze PCR samples for future use in N gene digestion analysis.

5) Agarose gel electrophoresis

- Make the following solutions if needed (5X TBE buffer – dissolve a sachet of Sigma 5X TBE in 1L of sterile distilled water and mix on the magnetic stirrer – dilute 10X prior to use; Loading dye – 3g Glycerol, 7ml sterile distilled water and a dash of bromophenol blue)
- Depending on the number of samples either a 30ml gel (14 samples and 2 ladders) or 100ml gel (26 samples and 2 ladders) should be made. For a 30ml gel, make up in a 250ml conical flask 0.3g agarose and 30ml 0.5X TBE.
- Plug the top of the flask loosely with paper roll and microwave for 1 minute at power level 9 to melt the agarose.
- Cool the base of the flask under cold running water (care should be taken not to allow any water to get in the flask, or to burn yourself) until it can be held in your hand without overheating.
- Add 3ul 10mg/ml ethidium bromide (NOTE: ethidium bromide is very toxic. Always wear gloves when handling it) and gently swirl to mix.
- Pour the gel into a mini tank containing metal blocks and combs and allow to set. Remove any bubbles by piercing with a pipette tip.
- Gently load 8ul of ladder (consists of 1ul ladder, 2ul loading dye and 7ul water) in the first well and then load 20ul (multiplex second round product) or 8ul (N gene PCR) of your samples in the remaining wells

(one sample per well) – NOTE: it's best to add 8ul of your samples to clean 0.5ml tubes and add the 2ul of loading dye to these, rather than adding the dye to your PCR product tubes

- Run the gel using a power pack (ask someone to show you how to use the particular pack you are going to use, as they all work differently) at 30mA for 30 minutes
- For a 100ml gel, use 1g agarose in the 100ml 0.5X TBE and microwave for 2 minutes and pour into a large gel tray sealed at each end with tape

6) Digestion of N gene PCR products for typing of the N gene
   - The following digest mixes are made in separate labelled 0.5ml tubes for each sample to be typed from the N gene PCR product to give us the NP patterns:

	Volume to add per sample				
Reagents	***Hind*III**	***Pst*I**	***Bgl*II**	***Hae*III**	***Rsa*I**
10X buffer (colour coded with enzyme cap	2ul	2ul	2ul	2ul	2ul
Enzyme	0.5ul	0.5ul	0.5ul	0.5ul	0.5ul
Sterile Distilled Water	10.5ul	10.5ul	10.5ul	10.5ul	10.5ul
Cleaned PCR product	7ul	7ul	7ul	7ul	7ul

- Once made up, these tubes are incubated in the 37°C waterbath for 1 hour and then run on a 2% agarose gel (for mini gels, use 30ml 0.5X TBE, 0.6g agarose) and the patterns analysed (see below for example digestion patterns)
- When running these digests on the gel, group by sample, so that on the gel you have sample1 (*Hind*III, *Pst*I, *Bgl*II, *Hae*III, *Rsa*I), then sample2 (*Hind*III, *Pst*I, *Bgl*II, *Hae*III, *Rsa*I), etc...

7) Patterns for known RSV types

- The following table shows which enzymes should cut the N gene PCR products for the commonly known NP types:

	Enzymes (cut = +; uncut = -)					A type or B type RSV
NP group	***Hind*III**	***Pst*I**	***Bgl*II**	***Hae*III**	***Rsa*I**	
NP1						B
NP2	-	-	-	+	+	A
NP3	-	-	+	-	+	B
NP4	-	-	+	+	+	A
NP5						
NP6						
NP7						
NP8						

8) G gene RT-PCR from random primer cDNA used in multiplex analysis

- Set up the primary PCR reactions in the clean room by mixing the following reagents per sample:

Reagent	Volume per sample
Taq Mastermix	25ul
25mM $MgCl_2$	2ul
AG20+F164 (A type RSV) *or* BG10+F164 (B type RSV) primers (10pmol/ul each, premixed)	2ul
Sterile Distilled Water	18ul

- Aliquot 46ul of the mix into individual labelled 0.5ml tubes.
- Take the PCR mixes to the main lab and add 4ul of each sample's random primer cDNA to the corresponding PCR tubes and load the tubes into the PCR block and run on programme 11 (same as above for N gene)
- Once complete, set up the secondary PCR reactions in the clean room as follows and freeze the remainder of the primary PCR products for future use:

Reagent	Volume per sample
Taq Mastermix	50ul
25mM $MgCl_2$	
10pmol/ul BG10+F1 primer mix	2ul
Sterile Distilled Water	44ul

- Aliquot 98ul of the mix into individual labelled 0.5ml tubes.
- Take the PCR mixes to the main lab and add 2ul of each primary PCR sample to the corresponding PCR tubes and load the tubes into the PCR block and run programme 11 (same as above for N gene)

- Proceed to agarose gel electrophoresis to view bands. Successful samples are subjected to PCR clean-up using the Qiagen PCR purification kit and stored at –20C.

8) Digestion of G gene RT-PCR products for typing

- For any G gene RT-PCR products that showed clear bands in the gel, clean up the whole product, using the Qiagen PCR purification kit.
- The following digest mixes were made in separate labelled 0.5ml tubes for each sample to be typed from the G gene PCR product to give us the G gene patterns:

	Volume to add per sample			
Reagents	***Alu*I**	***Taq*I**	***Mbo*I**	***Mse*I**
10X buffer (colour coded with enzyme cap	2ul	2ul	2ul	2ul
Enzyme	0.5ul	0.5ul	0.5ul	0.5ul
Sterile Distilled Water	10.5ul	10.5ul	10.5ul	10.5ul
Cleaned PCR product	7ul	7ul	7ul	7ul

- Once made up, these tubes were incubated in the 37°C waterbath for 1 hour (Except for *Taq*I which is incubated at 60°C for 1 hour) and then run on a 2% agarose gel (for mini gels, use 30ml 0.5X TBE, 0.6g agarose) and the patterns analysed (see below for example digestion patterns).
- Run these products in groups according to the enzyme used. For example if you have 5 samples that are digested with these enzymes, run all 5 *Alu*I digests together, then all *Taq*I digests together, then all *Mbo*I digests and finally all *Mse*I digests (i.e. group by enzyme)

Appendix M: Standard Operating Procedure for Oral-Fluid ELISA

IgG Capture Assay

Reagents:

- **COAT** (rabbit IgG in sodium carbonate buffer).
- **PBST** (phosphate buffered saline/0.05% Tween).
- **STANDARDS** (180µg/ml IgG calibrant standards (Binding Site, Edgbaston, England).
- **CONJUGATE** (Horseradish peroxidase-conjugated rabbit antihuman IgG (Dako)
- **SUBSTRATE** (*O*-phenylenediamine).
- **ACID** (2.5M sulphuric acid, H_2SO_4).

Reagent Preparation

- Sodium Carbonate Coating Buffer
  1.59g N_2CO_3 (4.29g if decahydrate) and 2.93g N_2HCO_3 in 1L distilled water

- Blocking Solution 5g of Marvel in 100ml PBST
- PBS-T (PBS with 0.05 % Tween 20)

Procedure

1. Frozen saliva samples thawed

2. Dilute rabbit anti-human IgG (Dako Ltd., Cambridge) 1/3000 in Sodium Carbonate ($NaCO_3$) coating buffer

3. Coat 96 well plates overnight at 37 °C with 100µl/well diluted rabbit anti-human IgG

4. Wash plates 4 time with PBS-T (PBS containing 0.05% Tween)

5. Block plates with 200µl/ well of 5% Marvel in PBS-T. Seal plates and incubate at 37 °C for 60 minutes

6. Flick off block and bang plate on towel

7. Meanwhile, prepare IgG calibrant standards; 0.0039-2.5 mg/l by doubling dilutions in PBS-T, assay in duplicate (Take 45µl of 180mg/l standard in 135µl of PBST to make a 45mg/l std. Take 10µl of the 45 mg/l in 170µl PBST to make the 2.5mg/l standard and double dilute to 0.0039 mg/l standard).

8. Dilute saliva 1/100 (or doubling dilutions 1/10 to 1/20480) in PBS-T and assay in triplicate alongside standards in duplicate (10µl sample or std /well in 40µl/ well blocking solution). Seal plates and incubate at 37 °C for 90 minutes.

9. Wash 8x with PBS-T. Flick off and bang plate dry on paper towel

10. Add 100µl conjugate (Horseradish Peroxidase-conjugated rabbit anti-human IgG (Dako) diluted 1/1000 and incubate at 37 °C for 2 hours

11. Wash 8 times with PBS-T. Flick off and bang plate dry on paper towel

12. Develop with substrate (*O*-phenylenediamine) for 10 minutes (30mg tablet in 30mls PBS, add 30µl H_2O_2 prepared just before use. Use 100µl per well.

13. Stop reaction with 50µl 2.5 mol/L H_2SO_4 and read at 492nm.

RSV specific ELISA

1. Block plates with 200µl/ well of 5% Marvel in PBS-T. Seal plates and incubate at at 37 °C for 60 minutes

2. Flick off block and bang plate on towel

3. Add saliva 50µl/ well (Diluted 1/2 to 1/64 in PBS-T or diluted to a 1mg/l concentration) and assay in triplicate. Seal plates and incubate at 37 °C for 90 minutes.

4. Wash 4x with PBS_T. Flick off and bang plate dry on paper towel

5. Add 100µl/ well conjugate (Horseradish peroxidase-conjugated rabbit/goat anti-human IgG (Dako)) diluted 1/1000 and incubate at 37 °C for 2 hours

6. Wash 8 times with PBS-T. Flick off and bang plate dry on paper towel

7. Develop with substrate (*O*-phenylenediamine) for 10 minutes (30mg tablet in 30mls PBS, add 30µl H_2O_2 prepared just before use).

8. Stop reaction with 50µl/ well of 2.5M H_2SO_4 per well and read at 492nm.

Appendix N: SOP for Cell Culture and Preparation of Lysate

Cell culture work to produce virus antigen and coated plates were supplied by Dr P. Scott from Health Protection Agency in Birmingham

CELL CULTURE

HEp-2

MATERIALS

500ml Minimal Essential Medium + Earles salts + L-Glutamine (MEM)
5ml HEPES buffer (1M)
5ml non-essential amino acids (NEAA, 100X)
5ml pen/strep (10 000U)
25ml Fetal calf serum (FCS, 5%)
Place CMEM (MEM + HEPES + NEAA + pen/strep + FCS) in water bath at 37°C to warm up.
Take 1 vial of HEp-2 cells from liquid N_2 quickly thaw a water bath at 37°C (NB. wear face protection when handling vials from liquid N_2).
Place cells in 5 ml of pre-warmed CMEM in a T-25 flask and allow cells to settle for 5-6 hours and then replace medium.

1. Incubate and check cells daily. When cells confluent, go on to split.

SPLITTING HEp-2 CELLS

MATERIALS

Trypsin-EDTA

CMEM

Sterile PBS

1. Pre-warm CMEM and trypsin-EDTA to 37°C prior to use.
2. Pipette off old media, add 5ml PBS to wash off medium (repeat).
3. Add 5ml trypsin-EDTA, wash cells briefly with this and then pipette off leaving just a film on the cells, incubate at 37°C but check every few minutes
4. When cells starting to lift off, tap flask firmly on your hand so all detach
5. Add 10 ml medium. Pipette up and down to break up cell clumps but don't make bubbles.
6. Distribute cells into new flasks and add appropriate amount of new medium.
7. Incubate all flasks at 37°C, check on flasks daily to monitor growth.

PREPARATION OF HEp-2 STOCK

1. Trypsinize cells, from a 150cm^2 T-flask and place cells in a centrifuge tube and spin at 1000rpm for 10mins.
2. Gently pipette off medium without disturbing pellet, and then gently re-suspend pellet in 2.4 ml of CMEM.

3. Label 6 ampoules (Name, date, passage #) and for each vial, add:

   0.4ml cells

   0.1ml DMSO

   0.5ml FCS

4. Do a step wise freeze down of cells (i.e. –70°C overnight and then transfer as quickly as possible to liquid N_2).
5. Take note of rack and tray used.

RESPIRATORY SYNCYTIAL VIRUS

MATERIALS
Slow Grow Medium

500ml Minimal Essential Medium + Earles salts + L-Glutamine (MEM)

5ml HEPES buffer (1M)

5ml non-essential amino acids (NEAA, 100X)

5ml pen/strep (10 000U)

10ml Fetal calf serum (FCS, 2%)

1. Place slow grow CMEM (MEM + HEPES + NEAA + pen/strep + FCS) in a water bath at 37°C to warm up.
2. Take 1 vial of A2 virus from -70 and thaw rapidly in 37°C waterbath.
3. Take 2 T-25 flasks of Hep-2 cells showing approximately 75% confluency and pipette off old media until about 1 ml left.
4. Infect 1 flask by adding A2 virus to T-25 flask, the other T-25 flask is a CONTROL flask.
5. Incubate flasks at 33°C for 2-3 hours, gently tipping both flasks every 15 mins, to allow for virus adherence to Hep-2 cells.
6. Top up the volume to bring up the final volume to 5ml.
7. Re-incubate both flasks at 33 °C and check daily; look for cytopathic effect (CPE) and change in pH of medium in infected flask in comparison to control.
8. Go on to make up RSV stocks.

PREPARATION OF VIRUS LYSATE for ELISA
MATERIALS
Nonidet-P40 (NP40)

PBS

Cell scraper

1. When extensive CPE (50-75%) is noted in infected cell, scrape cells into medium, and place in a centrifuge tube, similarly for control flasks.
2. Spin at 1500 rpm for 10 mins and then decant off supernatant (CAUTION: Infectious)
3. Resuspend pellet in 10ml PBS, spin at 1500 rpm for 10 mins.
4. Decant off supernatant (CAUTION: Infectious).
5. Resuspend in 10ml sterile water + 0.5% NP40 and vortex very hard.
6. Spin at 1500 rpm for 10 mins.
7. Aliquot supernatant in 2ml quantities and store at –70°C until further use.

PREPARATION OF VIRUS STOCK
MATERIALS

Sterile glass beads

1. When note extensive CPE (50-75%) harvest.
2. Label 7 ampoules appropriately (Name, date, virus stock no.)
3. Take sterile glass beads and place enough to cover the bottom surface of a clean bijou.
4. Pour these into the T-25 flask, vigourously shake to dislodge cells (check under light microscope). You may need to bkast cells off using pipette also.
5. Carefully pipette cell suspension, avoiding sucking up glass beads, and divide the cell suspension evenly between the 7 ampoules and store cells at –70°C until further use.

NB general: Bubbles kill cells but not virus! It is always best to thaw quickly not slowly. If you want to grow up a really good stock of virus in the early passes avoid freeze thawing, and instead inoculate a fresh flask directly with virus from the first flask.

ELISA

Plate preparation

Dilution of lysate must first be determined.

Dilute lysate 1:2, 4, 8, 16, 32, 64 in water + 0.5%NP40 (both infected and uninfected)

(To prepare plates for screening of sera then coat only with predetermined dilution of lysate)

Add 50 µl to each well (always do wells in triplicate) of 96 well TC plates

Allow to dry o/n in cat 1 hood

Fix with 100 µl/well 80% acetone in H20 for 10 mins.

Flick off acetone and allow to dry.

Store plates in bag at –20C

To carry out ELISA

Block plates with 200 µl/well 5% Marvel in PBS-T (PBS with 0.05% Tween 20). Seal plates and incubate for 1 hr at 37C. (Marvel is essential: some other brands don't work)

Flick off block and bang plate on paper towel

Meanwhile dilute sera in block (ie 5% Marvel in PBS-T)

Add sera 100 µl/well (For Mabs use 1/100 & 1/500 dilution; for sera 1/100 and doubling dilns to 1:3200)

Seal plate and incubate for 1.5 hr at 37C

If plate washer working, wash plates 4 times with 200 ul PBS-T.

If plate washer not working: remove sera and discard tips and contents as infectious. Add 200 µl PBS-T to each well. Flick off and bang plate dry on paper towel. Repeat 3 times

Add 100 µl secondary antibody to each well (for Mabs use Biorad goat anti mouse, HRP conjugated, for sera goat anti human HRP conjugated, both diluted 1:1000 in block)

Incubate at 37 °C for 1 hour

Wash as above

Develop with 100 µl O-phenylenediamine (10 mg tablet in 10 ml PBS, add 10 µl H2O2 (fresh) just before use); care: carcinogenic, per well for 10 mins

Add 50 µl 2.5 M H2SO4 per well. Read at 492 nm

For determination of optimum lysate dilution:

Dilute lysate in H2O with 0.5% NP40 as above.

Coat plate: A1-3,1: 2; A4-6, 1: 4 (infected) , A7-9, 1:2, A10-12 1:4 (uninfected)

B,1:8, 1:16

C, 1:32, 1:64

D: leave empty (Ab Control)

E-H repeat as above

Fix and block etc

Antibodies: 021/1G diluted 1:500 for A-D, 0121/2G dil 1:500 for E-H

You want to choose the best lysate dilution that gives a good positive signal with the infected but not with the uninfected lysate.

Layout of antigen dilution plate

1:2	1:2	1:2	1:4	1:4	1:4	1:2	1:2	1:2	1:4	1:4	1:4
1:8	1:8	1:8	1:16	1:16	1:16	1:8	1:8	1:8	1:16	1:16	1:16
1:32	1:32	1:32	1:64	1:64	1:64	1:32	1:32	1:32	1:64	1:64	1:64
No antigen						No mock antigen					
1:2	1:2	1:2	1:4	1:4	1:4	1:2	1:2	1:2	1:4	1:4	1:4
1:8	1:8	1:8	1:16	1:16	1:16	1:8	1:8	1:8	1:16	1:16	1:16
1:32	1:32	1:32	1:64	1:64	1:64	1:32	1:32	1:32	1:64	1:64	1:64
No antigen						No mock antigen					

021/1G

021/2G

A2 infected lysate

Mock control infected lysate

Layout of serum dilution plate

1:10	1:10	1:10	1:20	1:20	1:20	1:10	1:10	1:10	1:20	1:20	1:20
1:40	1:40	1:40	1:80	1:80	1:80	1:40	1:40	1:40	1:80	1:80	1:80
1:160	1:160	1:160	1:320	1:320	1:320	1:160	1:160	1:160	1:320	1:320	1:320
1:10 serum						1:10 serum					
1:10	1:10	1:10	1:20	1:20	1:20	1:10	1:10	1:10	1:20	1:20	1:20
1:40	1:40	1:40	1:80	1:80	1:80	1:40	1:40	1:40	1:80	1:80	1:80
1:160	1:160	1:160	1:320	1:320	1:320	1:160	1:160	1:160	1:320	1:320	1:320
1:10 serum						1:10 serum					

1:32 lysate

1:64 lysate

A2 infected lysate

Mock control infected lysate

Appendix O: *Calculating weights for observations within or between age classes by age strata*

Total incidence

	Observed			Expected		Weights	
age class	non-epidemic *pdo*	epidemic *pdo*	total *pdo*	non_epidemic *pdo*	epidemic *pdo*	non_epidemic	epidemic
0-11m	20214.6	7058.199	27273	18514	8759	**0.91587**	**1.24093**
12-17m	6847	5578	12425	8435	3990	**1.23188**	**0.71537**
18-24m	7092	6038	13130	8913	4217	**1.25680**	**0.69837**
25-36m	9792	3124	12916	8768	4148	**0.89542**	**1.32779**
37-59m	18662	11442	30104	20436	9668	**1.09506**	**0.84496**
5-9yr	26554	15168.2	41722	28323	13399	**1.06662**	**0.88338**
>9yr	23069	13197	36266	24619	11647	**1.06719**	**0.88255**
Total	112230.6 307	61605.4 169	173836 476	115334	58503	**1.02765**	**0.94963**

Formula	Observed			Expected		Weights	
age class, *i*	non-epidemic *pdo*	epidemic *pdo*	total *pdo*	non_epidemic *pdo*	epidemic *pdo*	non_epidemic	epidemic
0-11m	$Yn(1)$	$Ye(1)$	$Yn(1)+Ye(1)$				
12-17m	$Yn(2)$	$Ye(2)$	$Yn(2)+Ye(2)$				
18-24m	$Yn(3)$	$Ye(3)$	...				
25-36m	..	..	...				
37-59m	$Yn(i)$	$Ye(i)$	$\Sigma_k Y_k(i)$	$Pn.\Sigma_k Y_k(i)$	$Pe.\Sigma_k Y_k(i)$	$(Pn.\Sigma_k Y_k(i))/Yn(i)$	$(Pe.\Sigma_k Y_k(i))/Ye(i)$
5-9yr	...	...	...				
>9 yr	$Yn(7)$	$Ye(7)$	$\Sigma_k Y_k(7)$				
Total	$\Sigma_i(Y_n(i))$	$\Sigma_i(Y_e(i))$	$\Sigma_{ki}(Y_k(i))$				

$Y_k(i)$	observation period (days) for non-epidemic (k=n) or epidemic (k=e) for age class i =1,7		
Pe	expected proportion of year with epidemic	16.7/52	0.32115
Pn	expected proportion of year without epidemic	35.3/52	0.67885

Calculation of epidemic/non-epidemic portions of year

						weeks inclusive	
year	epidemic weeks	non-epidemic weeks		year	epidemic	start	finish
2002	20	32	52	2002	1	11	26
2003	15	37	52		2	49	52...
2004	28	24	52	2003	2	...1	15
2005	7	45	52	2004	3	2	22
sum	70	138	208		4	46	52...
wks/yr	16.7	35.3	52	2005	4	...1	7

Appendix P: Incidence estimated by age and child class
Incidence estimates of ARI by age and child class including 95% CI

	Number of events, person years of observation		
	(Incidence + 95% CI)		
Age (years)	Pre-school children with siblings in school	Pre-school children without siblings in school	School children
0	460, 56.9 8.06 (7.19, 9.03)	134, 18.6 7.23(5.10, 10.25)	
1	537, 52.4 10.22 (9.09, 11.49)	127, 15.4 8.09 (5.58, 11.73)	
2	161, 27 5.97 (5.03, 7.07)	41, 8.6 4.71 (2.74, 8.09)	
3	255, 35.3 6.73 (5.46, 8.30)	60, 10.7 5.43 (2.82, 10.46)	13, 1.30 8.24(2.60, 26.06)
4	121, 23.6 4.78 (3.56, 6.42)	33, 7.7 4.06 (1.63, 10.10)	32, 8.6 3.45 (1.39, 8.56)
5	36, 9.30 3.81 (2.46, 5.90)	22, 5.6 3.70 (1.18, 11.58)	48, 15 3.11 (1.14, 8.47)
6	12, 2.59 4.74 (1.87, 12.01)	12, 2.88 4.51 (0.51, 39.81)	79, 25.7 2.99 (0.44, 20.17)
7	0	1, 0.57 1.75 (0.21, 14.23)	78, 29.5 2.63 (0.04, 177.06)
8			59, 22.1 2.60 (1.58, 4.29)
9			26, 14.9 1.83 (0.98, 3.43)
10			31, 18.7 1.61 (1.02, 2.52)
11	0 2, 0.66 3.04 (0.53, 17.62)		19, 21.5 0.86 (0.01, 119.82)
12	1, 0.42 2.40 (0.25, 23.20)		27, 16.1 1.68 (0.03, 87.53)
13			35, 15.2 2.34 (1.56, 3.44)
14			17, 9.1 1.62 (0.76, 3.47)
15			5, 4.8 1.80 (0.49, 6.60)

Incidence estimates (95 % CI) of RSV infection by age and child class

	Incidence + 95% CI		
Age class	Pre-school children with siblings in school	Pre-school children without siblings in school	School children
0-11m	351 (227- 544)	431 (123-1517	
12-17m	741 (478-1,148)	509 (112-2,309)	
18-24m	533 (321-883)	361 (63-2,069)	
25-36m	377 (203-701)	111 (8-1,608)	0
37-59m	339 (216- 532)	291 (69-1,223)	99 (8-1,164)
5-9yr	165 (41-661)	333 (14-7,967)	86 (5-1,611)

Appendix Q: Incidence rates of RSV infection and disease stratified by age class

Table 1: Incidence rates of total infection and disease stratified by age class

	Incidence/1000 cyo (95% CI)		
Age class	Infection	RSV-LRTI	severe LRTI
0-11m	459 (316, 665)	33 (8, 131)	33 (8, 131)
12-17m	495 (197, 1,242)	165 (9, 3,114)	41 (1, 1,174)
18-24m	346 (132, 911)	58 (2, 1,385)	0
25-36m	409 (140, 1,194)	37 (1, 1,645)	0
37-59m	254 (102, 633)	0	0
60-108m	100 (36, 281)	0	0
109+m	9 (1, 95)	0	0

Table 2: Incidence of primary infection and disease stratified by age class

	Incidence/1000 cyo (95% CI)		
Age class	Infection	RSV-LRTI	severe LRTI
0-11m	406 (270, 612)	35 (9, 141)	35 (9,141)
12-17m	873 (315, 2,414)	349 (18, 6,609)	87 (3, 2,490)
18-24m	267 (72, 995)	45 (1, 1,969)	

Appendix R: Effect of weighting on observations using different Age classifications

Table 1: Calculating weights for observations within or between age classes by age strata

	Observed		Expected		Weights	
	non-epidemic	epidemic	non_epidemic	epidemic	non_epidemic	epidemic
age class	*pdo*	*pdo*	*pdo*	*pdo*		
0-11m	20215	7058	18514	8759	0.91587	1.24093
12-17m	6847	5578	8435	3990	1.23188	0.71537
18-24m	7092	6038	8913	4217	1.25680	0.69837
25-36m	9792	3124	8768	4148	0.89542	1.32779
37-59m	18662	11442	20436	9668	1.09506	0.84496
5-9yr	26554	15168	28323	13399	1.06662	0.88338
10+yr	23069	13197	24619	11647	1.06719	0.88255
Total	112231	61605	115334	58503	1.02765	0.94963

Table 2: Effect on weighting of combining age classes

	Observed		Expected		Weights	
	non-epidemic	epidemic	non_epidemic	epidemic	non_epidemic	epidemic
age class	*pdo*	*pdo*	*pdo*	*pdo*		
0-17m	27062	12636	26949	12749	0.99583	1.00893
18-36m	16884	9162	17681	8365	1.04722	0.91299
37-9yr	45216	26610	48759	23067	1.07836	0.86686
>9yr	23069	13197	24619	11647	1.06719	0.88255

Risk Factor Survey

Included in this section are the risk factor survey questionnaire and the accompanying rubric for its completion.

Appendix S: Risk Factors Survey Questionnaire

RSV RISK FACTORS STUDY QUESTIONNAIRE

1.General

Date today .....(DD/MM/YYYY)__ __ /__ __/ 2004

PID [ ][ ][ ][ ][ ][ ][ ] KBC # [ ][ ][ ][ ][ ] RSV # [ ][ ][ ]

Child names 1 _____ 2 _____ 3 _____

Date of birth .....(DD/MM/YYYY) __ __ /__ __/__ __

Birth weight .....Kg [ ].[ ][ ]

Sex: .....M/F [ ]

Number of babies born. *Circle one*

Singleton	Twins	Triplets	Other

If other, *specify* .....

2.Primary care taker details.

Names 1 _____ 2 _____ 3 _____

Age of PCT:..... years [ ][ ]

Sex of PCT:.....M/F [ ]

2.1 How many years of schooling/education have you had?.....number [ ][ ]

2.2 Can you read? *{Give reading card}*..... Y/N [ ]

2.3 Occupation of the major income provider. *Circle only one.*

Police/armed forces	Farmer	Hotel industry	Fishing	Petty Trader	Registered Business
Transport industry	Carpenter/ Mason	Health personnel	Teacher	Civil Servant	Other

If other; specify_____

2.4 Do you rely on assistance from relative/person(s) living away to make ends meet? .....Y/N [ ]

3. Demographic, Social-economic and Environmental factors

3.1 How many months of its life has the child breastfed? *(Duration of breastfeeding)*

Never	
All of life	
Part of life *(Indicate number of months)*	

3.2 At what age was the child weaned?.....age in months [ ]

3.3 How many people are in the family unit?..... number [ ][ ]

3.4 How many of these are children *{< 15 years}*?..... .....number [ ][ ]
(enter details of each in the table below)

{On this space make a diagram of the house and the sleeping arrangements in relation to the most recent epidemic} ***Applies to those in the household study ONLY***

If one or more child <15 years, enter details for each in the Table below (if >10 continue overleaf):

#	Name (First name + Initial only)	Age Yrs	Sex M/F	Live in same house as index Y/N			Sleep in same room as index Y/N			Sleep in same bed as index Y/N			Attended school Epidemic Y/N		
				Epi1	**Epi2**	**Epi3**	**Epi1**	**Epi2**	**Epi3**	**Epi1**	**Epi2**	**Epi3**	**Epi1**	**Epi2**	**Epi3**
1															
2															
3															
4															
5															
6															
7															
8															
9															
10															

Epidemic 1 (Mar 02-July 02) ***Epidemic 2 (Dec 02- April 03)*** ***Epidemic 3 (Jan 04-May 04)***

3.5 How many adults sleep in the same room as the index/infant?.....number [ ][ ]

3.6 How many siblings of the index have passed away? ..... number [ ]

3.7 Where does the index child normally remain during the day? *Circle one*

Inside the home	Outside home

3.8 Who usually takes care of the index child during the day? *Circle one*

Mother	Another family member	House help	Other

If other, *specify*_____

3.9 Which of the following do you have in the household at the present time?
Circle any applicable

Bicycle	Telephone/Cell	Radio	TV/video deck
Motorcycle	Electricity		

3.10 How many animals do you have Cows [ ] Goats [ ] Sheep [ ]

3.11 What type of main house does the PCT live in? *Circle one*

Mud wall house	Block wall house

3.12 Is this HH: *Circle one*

Owner occupied	Rented	No rent/with consent of owner	Other

If other, *specify* _____

3.13 What type of toilet do you have? *Circle one*

Flush toilet inside	Latrine inside	No toilet
Flush toilet outside	Latrine outside	

3.14 Do you burn your refuse within the homestead? ..... Y/N [ ]

3.15 What type of fuel do you use for cooking? *Circle more than one if necessary*

Fire wood	Wood Charcoal	Gas	Paraffin

3.16 Where do you do your cooking? *Circle more than one if necessary*

Same house as you sleep in	Different house from where you sleep	Outside

3.17 Where do you get water for domestic use? *Circle one*

Piped	Open or Closed Well	Surface

3.18 Site of piped or well water *Circle one*

Into dwelling	In compound or plot	Public

3.19 How many members of the family unit smoke?..... number [ ]

FIELD WORKER OBSERVATIONS - please write overleaf

Appendix T: Guide for completing RSV risk factor study questionnaire

Guide for completing RSV risk factor study questionnaire

Explain the purpose of the study to the head of the household/ mother before you begin

DATE.
Record date of interview. Other details concerning the child will be filled in automatically

1 Ask the Primary Caretaker (PCT) how many babies were born.

2.1 EDUCATION - Level of education attained. Record number of years of education. Put 0 if no school, 1 if attended only for one year etc, if completed secondary education and went to college, add years of college to years in school. [Some go to adult school - include this in years of schooling]

2.2 LITERACY - Ask if the PCT can read by letting them read the card provided. It is up to the fieldworker to classify the PCT as can read or not after hearing the PCT read this card. If they feel that the PCT can't read write 'N'

2.3 OCCUPATION - record a single occupation that nearest describes that of the major income earner of family unit. Draw big circle around number and occupation. If none apply, then circle 'other' and specify.

2.4 RELIANCE ON ASSISTANCE - Ask how often the family receives monetary assistance from another party (this may be a relative e.g. their children or from friends) and whether without this they would be unable to make ends meet. We are trying to establish the level of reliance on this help

3.1. BREASTFEEDING - Ask how many months the mother breastfed the child. Insert 0 for never, 99 if child still breast-feeding to date else fill in the actual length (completed months) of time the child was breast-fed.

3.2 EXCLUSIVITY OF BREAST FEEDING - Ask at what age the child was weaned

3.3. FAMILY UNIT - Record the number of people in this family unit - defined as the number that normally eat at the same meal. Give maximum inclusive of those who are temporarily away i.e. in boarding school

3.4 Record the number of children less than 15 years of age in the family unit.

Example diagram {*adult-filled triangle, children<15year-solid circle, index-blank cirle}*

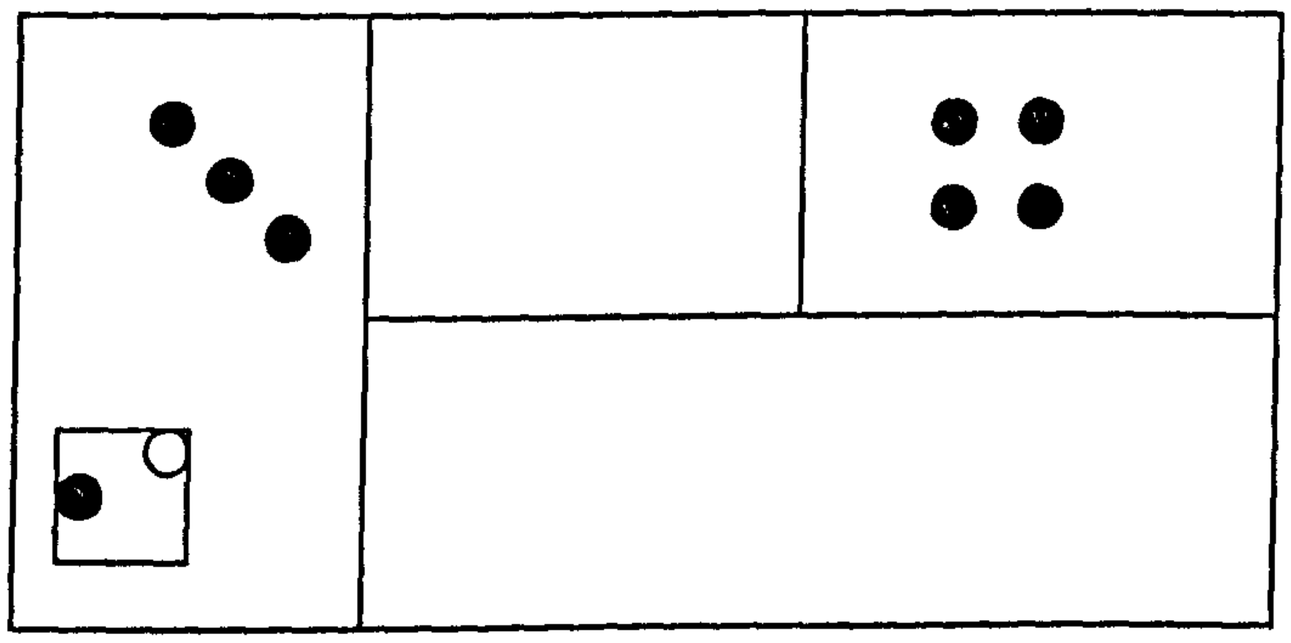


TABLE - Record the number of siblings for the index child (brothers/sisters).
Include half siblings or other children in the family unit who are not siblings.
- ask how many of the siblings are aged <5yrs , 5-14 yrs, 15+yrs (note explain grps)
- ask how many of the siblings are male

ASK THESE NEXT QUESTIONS IN RELATION TO THE THREE DIFFERENT EPIDEMICS AS SHOWN ON THE TABLE
- ask if the siblings live in the same house as the index child
- ask if the siblings sleep in the same room as the index child
- ask if the siblings normally attend school - primary and secondary
{Explain that some will have missed epidemic 1 and 2}

3.5 Ask how many adults sleep in the same room as index child

3.6 DEATHS - Ask if there is any brother or sister of index child who has passed away, then record the numbers who have died.

3.7. LOCATION DURING DAY - Ask PCT where the index child normally (PREDOMINANTLY i.e. more often than not) remains during the day. Circle appropriate location i.e. (1) Did the child normally remain in the home with mother or maid, or siblings in and around home (2) go out of the house for example with mother during her normal daily activities, to the work place, market, other compound/homestead

3.8 Ask with whom the child remains during the day. Other family member excludes the house help

3.9 OWNERSHIP LIST - Record what the index household/family unit owns which is a measure of social-economic standard. Circle as many boxes as are relevant.

3.10 Record count of how much livestock the household has.

3.11 WALL TYPE - record if wall of house of index child is concrete/stone/coral walled or mud walled.

3.12 RENTAL - Ask the PCT whether *they live in rented accommodation or not* (i.e. if they pay a rent to someone or what sort of arrangement they have).

3.13TOILET FACILITY - Record the type of toilet the family uses or no toilet, circle only one. And then record the location of the toilet.

3.14 BURNING RUBBISH - Record whether or not they burn their refuse within the house/homestead i.e. 'Y'or 'N'

3.15 FUEL - Ask for the kind of fuel they use for cooking in the house. Circle the MAIN fuel only.

3.16 COOKING LOCATION - Ask where the cooking actually takes place

3.17/8 WATER SUPPLY - Record water source for domestic use.

3.19 SMOKING - Ask how many members of the family unit smoke as a regular daily activity.

Appendix U: The means, ranges and distributions of variables used to compute the asset index

Variable	Mean	Std. Dev	Min	Max
Phone	0.2250	0.4179	0	1
Radio	0.6635	0.4730	0	1
Motor bike	0.0095	0.0969	0	1
Television set or video	0.1342	0.3412	0	1
Bicycle	0.4839	0.5002	0	1
Electricity	0.0737	0.2616	0	1
Cows	0.7599	2.1692	0	20
Goats	3.1550	9.8390	0	200
Sheep	0.3856	6.5931	0	150
Block walled house	0.3629	0.4813	0	1
Mud walled house	0.6371	0.4813	0	1
Owned house	0.7845	0.4116	0	1
Rented house	0.1853	0.3889	0	1
Not rented not owned	0.0302	0.1714	0	1
Flush toilet	0.0624	0.2421	0	1
Latrine	0.5936	0.4916	0	1
No toilet	0.3440	0.4755	0	1
Main fuel firewood	0.7448	0.4364	0	1
Main fuel charcoal	0.2420	0.4287	0	1
Main fuel non biomass	0.0302	0.1714	0	1
Drinking water from well	0.0586	0.2351	0	1
Drinking water from piped source	0.8960	0.3055	0	1
Water site owned	0.0832	0.2764	0	1
Water site public	0.7750	0.4179	0	1
Water site shared	0.1474	0.3549	0	1

Appendix V: WHO Child Growth Standards

Description

1. The macro (igrowup.ado) calculates z-scores for the five anthropometric indicators, weight-for-age, length/height-for-age, weight-for-length, weight-for-height and body mass index (BMI)-for-age, based on the WHO Child Growth Standards.

2. It flags any extreme (i.e. biologically implausible) z-scores for each indicator according to the following system:

Weight-for-age z-score (zwei) zwei < -6 or zwei > 5
Length/height-for-age z-score (zlen) zlen < -6 or zlen > 6
Weight-for-length/height z-score (zwfl) zwfl < -5 or zwfl > 5
BMI-for-age z-score (zbmi) zbmi < -5 or zbmi > 5

3. The macro produces sex- and age-specific estimates for the prevalence of under/over nutrition and summary statistics (mean and SD) of the z-scores for each indicator.

Pre-requisites

STATA Version 7.0 or higher is required to run "igrowup.ado". The macro requires a STATA data set containing age, sex and the anthropometric measurements. input variables are specified in the **Parameters** section.

Precautions:

1. Avoid any variable names starting with underscore "_" in the input STATA data set; otherwise they may be replaced by the derived ones created by the macro.

2. Avoid any temporary format names starting with underscore "_"; otherwise they may be replaced by the temporary ones created by the macro.

3. Avoid any STATA global macro variable names staring with underscore "_", except those defined by the system.

Macro Parameters

The macro requires 12 parameters that **must** be specified without any quotation marks.

- ***reflib***: to specify the package directory where the five STATA data sets containing the WHO Child Growth Standards are stored.

- ***datalib***: to specify the working directory where the input STATA data set containing anthropometric measurements is stored.

- ***datalab***: to specify the name that will prefix the output files (*datalab*_z, *datalab*_prev_st or *datalab*_prev_rc).

- *sex*: to specify the name of a variable containing sex information. If it is a numeric variable, its values must be, 1 for males and 2 for females. And if it is a character variable, it must be, "m" or "M" for males and "f" or "F" for females. Users must code any missing values as "." (for a numeric variable) or " " (for a character variable), in which case no z-scores will be calculated.

- *age*: to specify the name of a numeric variable containing age information. Age can be in either days or months. An exact age is expected here and should not be rounded if age is in months. Users must code any missing values as ".", in which case the age-related z-scores are not calculated.

- *ageunit*: to specify the age unit for the age variable. It must be specified as either "days" or "months" (they are *case sensitive*). To convert age in months to days, the macro multiplies by 30.4375 and rounds the result to integer for use with the reference tables.

- *weight*: to specify the name of a numeric variable containing body weight information. Body weight must be in kilograms. Users must code any missing values as ".", in which case weight-related z-scores are not calculated.

- *lenhei*: to specify the name of a numeric variable containing length (recumbent) or height (standing) information. Length or height must be in centimeters. Users must code any missing values as ".", in which case length- or height-related z-scores are not calculated. For children aged below 24 months (< 731 days) and measured standing, the macro converts the height to recumbent length by adding 0.7 cm; and for children aged 24 months or above who are measured in recumbent position, the macro converts the length to standing height by subtracting 0.7 cm. In other words, all the z-scores for children below 24 months are length-based, and height-based otherwise. The exported variable *_clenhei* is the converted length/ height according to age.

- *measure*: to specify the name of a character variable indicating whether recumbent length or standing height was measured. The values of this variable must be "l" or "L" for recumbent length, and "h" or "H" for standing height. Users must code any missing values as " ", and the macro imputes any missing values according to the following algorithm:
    a. If age is given, then it is recumbent length if the child's age is below 24 months (< 731 days), and standing height if the child's age is 24 months or above.
    b. If age is missing, then it is recumbent length if *lenhei* < 87 cm and standing height if *lenhei* ≥ 87 cm.
- *oedema*: to specify the name of the character variable containing oedema information. The values of this variable must be "n" or "N" for non-oedema, and "y" or "Y" for oedema. Users must code any missing values as " ", and the macro assumes that they are non-oedema. For oedema cases, weight-related z-scores (_zwei, _zwfl and _zbmi) are not calculated, but they are treated as being < -3 SD in their prevalence estimations.

- *sw*: to specify the name of a numeric variable containing sampling weight. If "**sw=1**" for all records, the un-weighted analysis is performed. If otherwise specified, negative values in sampling weight are not allowed and in this case no prevalence tables will be produced.

- ***flagsys***: to specify how to handle flagged z-scores in the summary report. Only values 0 or 1 can be attributed to this parameter; if "**flagsys=1**", records with at least one flagged z-score (flagged, true missing or with oedema) are excluded for the computation of all prevalences (restricted analysis). This option would generate the output *datlab*_**prev_rc.xls**. If "**flagsys=0**", all available (non-missing and non-flagged) z-score values are used for each indicator-specific prevalence estimation and oedema cases contribute as mentioned above (standard and recommended analysis). This option would generate the output *datalab*_**prev_st.xls**.

Exported files

1. The macro creates, in the working directory, a STATA data set. The name of the data set is *datalab*_z.dta (see the **Parameters** section). This data set retains all the records and variables from the input STATA data set and adds on the following 11 variables derived by the macro:

Variable name Variable label
_agedays calculated age in days for deriving z score
_clenhei converted length/height (cm) for deriving z score
_cbmi calculated bmi=weight / squared(_clenhei)
_zwei Weight-for-age z-score
_zlen Length/ height-for-age z-score
_zwfl Weight-for-length/height z-score
_zbmi BMI-for-age z-score
_fwei Flag for _zwei < -6 or _zwe i> 5
_flen Flag for _zlen < -6 or _zlen > 6
_fwfl Flag for _zwfl < -5 or _zwf l> 5
_fbmi Flag for _zbmi < -5 or _zbmi > 5

2. The macro creates, in the working directory, a data set in XLS format that has the same data structure as the one in STATA format. The name of the data set is *datalab*_z.xls.

3. The macro creates, in the working directory, a MS Excel file that contains the point estimates and 95% confidence intervals for the indicator prevalences and z-scores summary statistics. Only children aged below 61 completed months are included in the analysis. The age groups are: Total (0-60), 0-6, 7-11, 12-23, 24-47 and 48-60 completed months. Where age is missing, only weight-for-length/ height z-scores can be derived and these cases are included in the Total (0-60) age group. The name of this file is *datalab_prev_st* or *datalab_prev_rc*. The extension "_prev_st" or "_prev_rc" depends on the type of analysis requested via the macro parameter ***flagsys*** (see the **Parameters** section).

Recommended setup and run

Step 1. Create a sub-directory, for example "D:\WHO igrowup STATA", where you wish to save the package (**igrowup_stata.zip**). This directory should be reserved only for the references tables (*anthro.dta) and the macro (igrowup.ado) that are contained in the zip file.

Step 2. Create a sub-directory, for example "D:\WHO igrowup workdata", where the example data (survey.dta) and your STATA input data can be stored and where all the macro output files will be written to.

Step 3. It is recommended that you start by loading and running the example code below (also found in survey.do) in the STATA do-file editor to see how the data should be prepared and to fill in the macro parameters according to their requirements.

```
/* Example: survey.do using survey.dta */
clear
set more 1
/* Higher memory might be necessary for larger datasets */
set memory 10m
/* Indicate to the Stata compiler where the igrowup.ado file is
stored*/
adopath + "D:\WHO igrowup STATA/"
/* Load the data file */
use "D:\WHO igrowup workdata\survey.dta", clear
/* generate the first three parameters reflib, datalib & datalab */
gen str60 reflib="D:\WHO igrowup STATA"
lab var reflib "Directory of reference tables"
gen str60 datalib="D:\WHO igrowup workdata"
lab var datalib "Directory for datafiles"
gen str30 datalab="mysurvey"
lab var datalab "Working file"
/* check the variable for "sex" 1 = male, 2=female */
desc gender
tab gender
/* check the variable for "age" */
desc agemons
summ agemons
/* define your ageunit */
gen str6 ageunit="months" /* or [gen ageunit="days"] */
lab var ageunit "=days or =months"
/* check the variable for body "weight" which must be in kgs*/
desc weight
summ weight
/* check the variable for "lenhei" which must be in cms */
desc height
summ height
/* check the variable for "measure"*/
/* NOTE: if not available, please create as [gen str1 measure=" "]*/
desc measure
tab measure
/* check the variable for "oedema"*/
/* NOTE: if not available, please create as [gen str1 oedema="n"]*/
desc oedema
tab oedema
```

```
/* check the variable for "sw" for the sampling weight */
/* NOTE: if not available, please create as [gen sw=1] */
desc sw
summ sw
/* define the variable "flagsys"=0 if standard tables;
"flagsys"= 1 if restricted tables */
gen flagsys=0 /* if standard table */
lab var flagsys "=0 Standard table;=1 Restricted table"
*replace flagsys=1 /* if restricted table */
/* Fill in the macro parameters to run the command */
igrowup reflib datalib datalab gender agemons ageunit weight height
measure oedema sw flagsys
```

References

WHO Multicentre Growth Reference Study Group. WHO Child Growth Standards: Length/height-for-age, weight-for-age, weight-for-length, weight-for-height and body mass index-for-age: Methods and development. Geneva: World Health Organization, 2006.

```
/*     eo_survey.do using survey.dta */

clear
set more 1

/* Higher memory might be necessary for larger datasets */

set memory 10m

/* Indicate to the Stata compiler where the igrowup.ado file is
stored*/
adopath + "C:\My Documents\EOkiro\WHOigrowupSTATA/"

/* Load the data file */

use "C:\My Documents\EOkiro\WHOigrowup workdata\survey.dta", clear

/* generate the first three parameters reflib, datalib & datalab  */

gen str60 reflib="C:\My Documents\EOkiro\WHOigrowupSTATA"
lab var reflib "Directory of reference tables"

gen str60 datalib="C:\My Documents\EOkiro\WHOigrowup workdata"
lab var datalib "Directory for datafiles"

gen str30 datalab="mysurvey"
lab var datalab "Working file"

/*     check the variable for "sex"  1 = male, 2=female */
desc gender
tab gender


/*     check the variable for "age"  */
desc agemons
summ agemons


/*     define your ageunit      */
gen str6 ageunit="months"
lab var ageunit "=months"


/*     check the variable for body "weight" which must be in kgs*/
desc weight
summ weight

/*     check the variable for "height" which must be in cms  */
desc height
summ height


/*     check the variable for "measure"*/
/*     NOTE: if not available, please create as [gen str1 measure="
"]*/
desc measure
tab measure
```

```
/*     check the variable for "oedema"*/
/*     NOTE: if not available, please create as [gen str1
oedema="n"]*/
desc oedema
tab oedema


/*     check the variable for "sw" for the sampling weight*/
/*     NOTE: if not available, please create as [gen sw=1]*/
desc sw
summ sw
```

```
/*     define the variable "flagsys"=0 if standard tables;
       "flagsys"= 1 if restricted tables                          */

gen flagsys=0              /* if standard table */
lab var flagsys "=0 Standard table;=1 Restricted table"
*replace flagsys=1         /* if restricted table */


/*     Fill in the macro parameters to run the command */
igrowup reflib datalib datalab gender agemons ageunit weight height
measure oedema sw flagsys
```

Appendix W: Results from univariate analysis

In this appendix we present results of the univariate risk factor analysis of the three outcomes studies RSVI, RSV-LRTI and all cause LRTI

Table 1: Univariate analyses of factors associated with LRTI using Cox Regression

Appendix W: Results from univariate analysis

In this appendix we present results of the univariate risk factor analysis of the three outcomes studies RSVI, RSV-LRTI and all cause LRTI

Table 1: Univariate analyses of factors associated with LRTI using Cox Regression

Potential risk factors	*Categories (number of events)*	*Hazard Ratio*	*P>|z|§*	*95% Conf. Interval*	
SES	poor (312)	-	-	-	-
SES	Middle (317)	1.02	0.896	0.81	1.28
SES	**rich (228)**	**0.70**	**0.003**	**0.55**	**0.89**
Gender	Female (429)	-	-	-	-
Gender	Male (428)	1.11	0.275	0.92	1.35
current age_0	0-5 months (245)	-	-	-	-
current age_1	**6-11 months (120)**	**0.58**	**0.000**	**0.43**	**0.77**
current age_3	**12- 17 months (290)**	**1.85**	**0.000**	**1.40**	**2.43**
current age_4	18+ months (202)	0.85	0.360	0.59	1.21
birth weight_0	1.2- <2.0 kg (40)	-	-	-	-
birth weight _1	2- <2.5 kg (76)	0.74	0.286	0.43	1.28
birth weight _2	2.5- <3.0 kg (276)	0.83	0.480	0.49	1.40
birth weight _3	3.0+ kg (447)	0.87	0.611	0.52	1.47
multiple babies_0	1 child (790)	-	-	-	-
multiple babies_1	Twins (50)	1.02	0.973	0.75	1.34
multiple babies_2	**Triplets (17)**	**3.85**	**0.000**	**3.35**	**4.42**
PCT*_agegroup_0	13-20 years (125)	-	-	-	-
PCT_agegroup_1	21-30 years (428)	0.99	0.980	0.74	1.31
PCT_agegroup_2	31-40 years (226)	0.93	0.654	0.68	1.28
PCT_agegroup_3	41-50 years (51)	0.88	0.537	0.58	1.32
PCT_agegrp_4	51-63 years (27)	1.36	0.083	0.96	1.92
education_0	no schooling (249)	-	-	-	-
education _1	1-7 years (353)	0.94	0.493	0.80	1.11
education _2	**8-12 years (241)**	**0.80**	**0.080**	**0.63**	**1.03**
***education* _3**	**>12 years (14)**	**0.34**	**0.000**	**0.21**	**0.55**
literacy	No (349)	-	-	-	-
literacy	**Yes (508)**	**0.82**	**0.038**	**0.67**	**0.98**
family_assisted	No (665)	-	-	-	-
family_assisted	Yes (192)	1.20	0.091	0.97	1.49
breast feeding _0	0-12 months (123)	-	-	-	-
breast feeding _1	13- 18 months (193)	1.01	0.899	0.81	1.27

breast feeding _2	*19- 23 months (81)*	*1.04*	*0.783*	*0.79*	*1.38*
breast feeding _3	24+ (460)	1.00	0.966	0.82	1.23
age at weaning _0	0-2 months (310)	-	-	-	-
age at weaning _1	3- 6 months (532)	0.84	0.085	0.69	1.02
age at weaning _2	7+ months (15)	0.59	0.166	0.28	1.24
Weight-age-z score_0	>-1 (507)	-	-	-	-
Weight-age-z score_1	-1.99 to -1 (205)	1.06	0.471	0.90	1.25
Weight-age-z score_2	**≤ -2 (145)**	**1.30**	**0.006**	**1.08**	**1.56**
Height-age-z score_0	>-1 (363)	-	-	-	-
Height-age-z score_1	**-1.99 to -1 (239)**	**1.32**	**0.001**	**1.12**	**1.56**
Height-age-z score_2	**≤ -2 (255)**	**1.60**	**0.000**	**1.36**	**1.89**
Weight-height-z score_0	>-1 (652)	-	-	-	-
Weight-height-z score_1	-1.99 to -1 (136)	0.91	0.327	0.76	1.10
Weight-height-z score_2	**≤ -2 (69)**	**1.41**	**0.007**	**1.10**	**1.81**
family_unit_0	2-5 members (261)	-	-	-	-
family_unit_1	6-7 members (203)	1.05	0.716	0.80	1.38
family_unit_2	8-10 members (207)	1.21	0.159	0.93	1.57
family_unit_3	11+ (186)	1.19	0.164	0.93	1.53
family_children_0	1-4 children (598)	-	-	-	-
family_children_1	6-10 children (214)	1.08	0.497	0.86	1.37
family_children_2	**11+children (45)**	**1.80**	**0.000**	**1.31**	**2.47**
siblings under 6 years_0	no siblings < 6 yrs (185)	-	-	-	-
siblings under 6 years_1	**1-2 siblings<6 yrs (517)**	**1.37**	**0.000**	**1.16**	**1.62**
siblings under 6 years_2	**3-4 siblings<6 yrs (129)**	**1.50**	**0.000**	**1.21**	**1.88**
siblings under 6 years_3	**5+ siblings<6 yrs (26)**	**1.70**	**0.000**	**1.28**	**2.26**
siblings 6 years or more_0	1-3 siblings 6+ yrs (290)	-	-	-	-
siblings 6 years or more_1	4-6 siblings 6+ yrs (292)	0.98	0.864	0.76	1.15
siblings 6 years or more_2	7-10 siblings 6+ yrs (202)	1.14	0.282	0.90	1.37
siblings 6 years or more_3	11+ siblings 6+ yrs (73)	1.00	0.994	0.72	1.29
male_siblings_0	0-2 male siblings (601)	-	-	-	-

male_siblings_1	**3-6 male siblings (241)**	**1.22**	**0.087**	**0.97**	**1.54**
male_siblings_2	7+ male siblings (15)	1.15	0.621	0.65	2.04
female_siblings_0	1-4 female sibs (426)	-	-	-	-
female_siblings_1	5-9 female sibs (343)	1.12	0.283	0.91	1.39
female_siblings_2	10+ female sibs (88)	1.01	0.935	0.75	1.36
# of dead siblings	None (640)	-	-	-	-
# of dead siblings	1 (157)	1.03	0.835	0.79	1.36
# of dead siblings	2 (60)	1.11	0.582	0.76	1.63
child stays	mostly inside house (817)	-	-	-	-
child stays	mostly outside house (40)	0.79	0.261	0.53	1.19
child's care_0	mother (670)	-	-	-	-
child's care _1	another family member (129)	1.24	0.060	0.99	1.56
child's care _2	**House help (20)**	**0.56**	**0.008**	**0.36**	**0.86**
child's care _3	School (0)	-	-	-	-
child's care _4	**mother / family member (38)**	**1.89**	**0.004**	**1.22**	**2.93**
house type	block walled (267)	-	-	-	-
house type	**mud walled (590)**	**1.37**	**0.002**	**1.12**	**1.67**
house ownership	owner occupied (690)	-	-	-	-
house ownership	rented (149)	0.87	0.316	0.67	1.14
house ownership	not owned/ rented (18)	0.62	0.104	0.35	1.10
toilet_type 0	no toilet (318)	-	-	-	-
toilet_type 1	**Flush (25)**	**0.40**	**0.001**	**0.24**	**0.67**
toilet_type_2	Latrine (514)	0.93	0.510	0.76	1.14
burn refuse	No (189)	-	-	-	-
burn refuse	yes (668)	1.16	0.172	0.94	1.45
main_fuel_0	gas/paraffin (16)	-	-	-	-
main_fuel_1	Charcoal (158)	1.61	0.070	0.96	2.68
main_fuel_2	**Firewood (660)**	**2.16**	**0.001**	**1.35**	**3.46**
main_fuel_3	**firewood/charcoal (23)**	**2.27**	**0.027**	**1.10**	**4.69**
cooking location_0	Outside (67)	-	-	-	-
cooking location_1	different house from sleep area (496)	1.01	0.945	0.66	1.55
cooking location_2	same house as you sleep (294)	0.92	0.690	0.59	1.42

water source_0	*both well & piped (43)*	-	-	-	-
water source_1	open/closed well (50)	0.95	0.837	0.57	1.57
water source_2	Piped (764)	0.86	0.431	0.60	1.24
watersite_0	own source (38)	-	-	-	-
watersite_1	**Public (686)**	**2.12**	**0.000**	**1.50**	**3.00**
watersite_2	**Shared (133)**	**2.07**	**0.001**	**1.37**	**3.11**
job description of major income provider_0	non-skilled (359)	-	-	-	-
job description of major income provider_1	Trade (146)	0.94	0.667	0.71	1.24
job description of major income provider_2	Skilled (384)	0.86	0.190	0.68	1.07
job description of major income provider_3	**Professional (68)**	**0.69**	**0.009**	**0.53**	**0.91**
smokers in HH_0	None (597)	-	-	-	-
smokers in HH_1	1 (216)	1.08	0.490	0.87	1.35
smokers in HH_2	2 or more (44)	1.15	0.442	0.81	1.64
adults in index's room_0	None (32)	-	-	-	-
adults in index's room_1	1 adult (263)	0.96	0.889	0.54	1.71
adults in index's room_2	2 adults (528)	0.79	0.411	0.45	1.38
adults in index's room_3	3+ adults (34)	0.85	0.666	0.41	1.78
# siblings <6yr living in same house as index_0	no siblings < 6 yrs (314)	-	-	-	-
# siblings <6yr living in same house as index_1	1 sibling (366)	1.10	0.384	0.89	1.35
# siblings <6yr living in same house as index_2	**2-4 siblings (150)**	**1.41**	**0.017**	**1.06**	**1.86**
# siblings 6+ years living in same house as index_0	no siblings 6+ years (296)	-	-	-	-
# siblings 6+ years living in same house as index_1	1-2 siblings (379)	1.14	0.238	0.92	1.42
# siblings 6+ years living in same house as index_2	3-4 siblings (117)	0.86	0.251	0.66	1.12
# siblings 6+ years living in same house as index_3	5-7 siblings (38)	1.45	0.128	0.90	2.36
# siblings <6yr sleeping in same room as index_0	no siblings < 6 yrs (411)	-	-	-	-
# siblings <6yr sleeping in same room as index_1	**1 -2 sibling < 6 yrs (408)**	**1.30**	**0.007**	**1.07**	**1.57**
# siblings <6yr sleeping in same room as index_2	3 siblings < 6 yrs (11)	1.41	0.541	0.47	4.28

# siblings 6+ years sleeping in same room as index_0	no siblings 6+ years (535)	-	-	-	-
# siblings 6+ years sleeping in same room as index_1	**1 siblings (194)**	**1.27**	**0.047**	**1.00**	**1.62**
# siblings 6+ years sleeping in same room as index_2	2-3 siblings (88)	0.95	0.741	0.72	1.26
# siblings 6+ years sleeping in same room as index_2	4 siblings (13)	0.83	0.631	0.38	1.79
# siblings <6yr sleeping in same bed as index_0	no siblings < 6 yrs (526)	-	-	-	-
# siblings <6yr sleeping in same bed as index_1	1 sibling (280)	1.22	0.061	0.99	1.51
# siblings <6yr sleeping in same bed as index_2	2-3 siblings (24)	1.48	0.222	0.79	2.75
# siblings 6+ years sleeping in same room as index_0	no siblings < 6 yrs (727)	-	-	-	-
# siblings 6+ years sleeping in same room as index_1	1-2 siblings (103)	0.99	0.919	0.75	1.29
# siblings <6yr going to school_0	no siblings < 6 yrs going to school (709)	-	-	-	-
# siblings <6yr going to school_1	**1 sibling (86)**	**0.77**	**0.033**	**0.60**	**0.98**
# siblings <6yr going to school_2	**2-3 siblings (35)**	**1.69**	**0.042**	**1.02**	**2.80**
# siblings 6+ years going to school_0	no siblings < 6 yrs going to school (235)	-	-	-	-
# siblings 6+ years going to school_1	1-2 sibling (347)	1.13	0.309	0.87	1.44
# siblings 6+ years going to school_2	3-5 siblings (216)	1.12	0.404	0.84	1.48
# siblings 6+ years going to school_3	6-9 siblings (32)	1.46	0.092	0.98	2.26

*PCT (Primary care taker)

§p-values (bold<0.05) for the tests of the null hypothesis that each hazard ratio is one.

Table 2: Results from univariate analysis using Cox regression with RSV-LRTI as the outcome variable. Significant variables are in bold.

Potential risk factors	*Categories (number of events)*	*Hazard Ratio*	*P>|z|*	*95% Conf. Interval*	
SES	Poor (28)	-	-	-	-
SES	Middle class (41)	1.49	0.089	0.94	2.37
SES	Rich (23)	0.80	0.401	0.48	1.34

Gender	Female (48)	-	-	-	-
Gender	Male (44)	1.03	0.883	0.70	1.52
current age_0	0-5 months (28)	-	-	-	-
current age_1	6-11 months (21)	1.20	0.753	0.38	3.77
current age_2	12- 17 months (16)	1.44	0.591	0.38	5.39
current age_3	18+ months (27)	1.06	0.927	0.28	4.05
birth weight_0	1.2- <2.0 kg (5)	-	-	-	-
birth weight _1	2- <2.5 kg (8)	0.66	0.424	0.24	1.82
birth weight _2	2.5- <3.0 kg (25)	0.68	0.370	0.29	1.59
birth weight _3	3.0+ kg (47)	0.78	0.552	0.34	1.78
multiple babies_0	1 child (83)	-	-	-	-
multiple babies_1	Twins (8)	1.54	0.203	0.79	3.01
multiple babies_2	Triplets (1)	1.99	0.351	0.47	8.50
PCT*_agegroup_0	13-22 years (6)	-	-	-	-
PCT_agegroup_1	23-30 years (44)	2.15	0.109	0.84	5.47
PCT_agegroup_2	31-40 years (28)	2.43	0.067	0.94	6.27
PCT_agegroup_3	**41-63 years (14)**	**3.65**	**0.012**	**1.34**	**9.98**
education_0	no schooling (25)	-	-	-	-
education _1	1-7 years (31)	0.84	0.530	0.49	1.44
education _2	8-12 years (34)	1.14	0.624	0.68	1.92
education _3	>12 years (2)	0.51	0.340	0.13	2.04
literate PCT	No (34)	-	-	-	-
literate PCT	Yes (58)	0.96	0.849	0.63	1.46
Family assisted	No (69)	-	-	-	-
Family assisted	Yes (23)	1.40	0.131	0.91	2.15
breast feeding _0	0- 12 months (13)	-	-	-	-
breast feeding _1	13- 18 months (21)	1.05	0.899	0.53	2.07
breast feeding _2	19- 23 months (11)	1.29	0.510	0.60	2.78
breast feeding _3	24+ (47)	1.02	0.929	0.57	1.84
age at weaning _0	0-2 months (30)	-	-	-	-
age at weaning _1	3- 6 months (60)	0.98	0.919	0.64	1.49
age at weaning _2	7+ months (2)	0.84	0.804	0.22	3.23
Weight-age-z score_0	>-1 (49)	-	-	-	-
Weight-age-z score_1	-1.99 to -1 (27)	1.37	0.179	0.86	2.17
Weight-age-z score_2	≤ -2 (16)	1.37	0.228	0.82	2.28
Height-age-z score_0	>-1 (38)	-	-	-	-
Height-age-z score_1	-1.99 to -1 (26)	1.43	0.162	0.87	2.36

Height-age-z score_2	**≤ -2 (28)**	**1.85**	**0.011**	**1.15**	**2.97**
Weight-height-z score_0	>-1 (71)	-	-	-	-
Weight-height-z score_1	-1.99 to -1 (13)	0.70	0.241	0.40	1.26
Weight-height-z score_2	≤ -2 (8)	1.17	0.656	0.58	2.37
family_unit_0	2-5 members (28)	-	-	-	-
family_unit_1	6-7 members (14)	0.66	0.208	0.35	1.26
family_unit_2	8-10 members (25)	1.35	0.248	0.81	2.27
family_unit_3	11+members (25)	1.50	0.116	0.90	2.48
family_children_0	1-5 children (60)	-	-	-	-
family_children_1	6-10 children (26)	1.32	0.214	0.85	2.04
family_children_2	**11+ children (6)**	**2.51**	**0.005**	**1.32**	**4.76**
siblings under 6 years_0	no siblings< 6 yrs (16)	-	-	-	-
siblings under 6 years_1	**1-2 siblings<6 yrs (58)**	**1.78**	**0.028**	**1.06**	**2.98**
siblings under 6 years_2	**3-4 siblings<6 yrs (15)**	**2.00**	**0.048**	**1.00**	**3.97**
siblings under 6 years_3	5+ siblings<6 yrs (3)	2.39	0.116	0.81	7.09
siblings 6 years or more_0	1-3 siblings 6+ yrs(30)	-	-	-	-
siblings 6 years or more_1	4-6 siblings 6+ yrs (26)	0.84	0.523	0.50	1.42
siblings 6 years or more_2	7-10 siblings 6+ yrs (26)	1.43	0.157	0.87	2.34
siblings 6 years or more_3	11+ siblings 6+ yrs (10)	1.36	0.350	0.71	2.60
male_siblings_0	0-2 male siblings (61)	-	-	-	-
male_siblings_1	3-6 male siblings (27)	1.36	0.170	0.87	2.10
male_siblings_2	**7+ male siblings (4)**	**3.28**	**0.000**	**1.76**	**6.17**
female_siblings_0	1-4 female sibs (39)	-	-	-	-
female_siblings_1	5-9 female sibs (42)	1.49	0.065	0.98	2.27
female_siblings_2	10+ female sibs (11)	1.40	0.272	0.77	2.56
# of dead siblings	None (67)	-	-	-	-
# of dead siblings	1 (18)	1.13	0.654	0.67	1.90
# of dead siblings	2 (7)	1.21	0.596	0.60	2.41
child stays	mostly inside house (91)	-	-	-	-
child stays	mostly outside house (1)	0.18	0.081	0.03	1.23
child's care_0	mother (73)	-	-	-	-
child's care _1	another family member (13)	1.12	0.692	0.63	1.98

child's care _2	house help (3)	0.78	0.658	0.27	2.31
child's care _3	school (0)	-	-	-	-
child's care _ 4	mother / family member (3)	1.28	0.622	0.48	3.38
house type	block walled (32)	-	-	-	-
house type	mud walled (60)	1.15	0.507	0.76	1.76
house ownership	owner occupied (80)	-	-	-	-
house ownership	**rented (10)**	**0.52**	**0.035**	**0.28**	**0.96**
house ownership	not owned/ rented (2)	0.55	0.370	0.15	2.00
toilet_type 0	no toilet (33)	-	-	-	-.
toilet_type 1	flush (1)	0.16	0.066	0.02	1.13
toilet_type_2	latrine (58)	1.03	0.889	0.68	1.57
burn refuse	No (18)	-	-	-	-
burn refuse	yes (74)	1.33	0.237	0.83	2.14
main_fuel_0	Charcoal (12)	-	-	-	-
main_fuel_1	gas/paraffin (0)	-	-	-	-
main_fuel_2	**firewood (79)**	**2.08**	**0.011**	**1.18**	**3.66**
main_fuel_3	firewood/charcoal (1)	0.82	0.831	0.13	5.29
cooking location_0	Outside (8)	-	-	-	-
cooking location_1	different house from sleep area (59)	1.00	0.994	0.52	1.92
cooking location_2	same house as you sleep (25)	0.65	0.228	0.32	1.31
water source_0	both well & piped (6)	-	-	-	-
water source_1	open/closed well (4)	0.52	0.279	0.16	1.70
water source_2	Piped (82)	0.66	0.270	0.31	1.38
watersite_0	Public (79)	-	-	-	-
watersite_1	own source (2)	**0.22**	**0.030**	**0.06**	**0.86**
watersite_2	Shared (11)	0.71	0.269	0.38	1.31
job description of MIP_0	non-skilled (35)	-	-	-	-
job description of MIP_1	Trade (21)	1.42	0.187	0.84	2.40
job description of MIP_2	Skilled (27)	0.84	0.470	0.53	1.34
job description of MIP_3	Professional (9)	0.97	0.936	0.46	2.03
smokers in HH_0	None (66)	-	-	-	-
smokers in HH_1	1 (21)	0.93	0.770	0.56	1.53
smokers in HH_2	2 or more (5)	1.21	0.687	0.47	3.11
adults in index's room_0	0 -1 adult (36)	-	-	-	-
adults in index's room_2	**2 or more adults (52)**	**0.66**	**0.047**	**0.44**	**1.00**

adults in index's room_3	3 or more adults (4)	0.82	0.653	0.34	1.95
# siblings <6yr living in same house as index_0	no siblings < 6 yrs (34)	-	-	-	-
# siblings <6yr living in same house as index_1	1 sibling (37)	1.00	0.995	0.65	1.55
# siblings <6yr living in same house as index_2	**2-4 siblings (21)**	**1.73**	**0.041**	**1.02**	**2.93**
# siblings 6+ years living in same house as index_0	no siblings 6+ years (34)	-	-	-	-
# siblings 6+ years living in same house as index_1	1-2 siblings (36)	0.92	0.726	0.59	1.44
# siblings 6+ years living in same house as index_2	3-4 siblings (15)	0.91	0.771	0.49	1.69
# siblings 6+ years living in same house as index_3	**5-7 siblings (7)**	**2.10**	**0.029**	**1.08**	**4.10**
# siblings <6yr sleeping in same room as index_0	no siblings < 6 yrs (50)	-	-	-	-
# siblings <6yr sleeping in same room as index_1	1 -3 sibling < 6 yrs (42)	1.05	0.797	0.71	1.55
# siblings 6+ years sleeping in same room as index_0	no siblings 6+ years (63)	-	-	-	-
# siblings 6+ years sleeping in same room as index_1	1 siblings (19)	1.05	0.855	0.64	1.71
# siblings 6+ years sleeping in same room as index_2	2-4 siblings (10)	0.77	0.462	0.39	1.53
# siblings <6yr sleeping in same bed as index_0	no siblings < 6 yrs (59)	-	-	-	-
# siblings <6yr sleeping in same bed as index_1	1-3 siblings (33)	1.16	0.456	0.78	1.72
# siblings 6+ years sleeping in same bed as index_0	no siblings < 6 yrs (84)	-	-	-	-
# siblings 6+ years sleeping in same bed as index_1	1-2 siblings (8)	0.64	0.186	0.33	1.24
# siblings <6yr going to school_0	no siblings < 6 yrs going to school (77)	-	-	-	-
# siblings <6yr going to school_1	1 sibling (10)	0.75	0.366	0.40	1.41
# siblings <6yr going to school_2	2-3 siblings (5)	1.77	0.182	0.76	4.12
# siblings 6+ years going to school_0	no siblings < 6 yrs going to school (25)	-	-	-	-
# siblings 6+ years going to school_1	1- 2 sibling(33)	1.02	0.944	0.61	1.70
# siblings 6+ years going to school_2	3-5 siblings (27)	1.31	0.310	0.78	2.20

# siblings 6+ years going to school 3	**6-9 siblings (7)**	**3.04**	**0.002**	**1.48**	**6.22**

*PCT (Primary care taker)
§p-values (bold<0.05) for the tests of the null hypothesis that each hazard ratio is one.

Table 3: Results form univariate analysis using Cox Regression with RSV infection as outcome

Potential risk factors	*Categories (number of events)*	*Hazard Ratio*	*P>|z|*	*95% Conf. Interval*	
SES	Poor (128)	-	-	-	-
SES	Middle class (134)	1.02	0.818	0.83	1.26
SES	**Rich (100)**	**0.75**	**0.015**	**0.60**	**0.95**
current age_0	0-5 months (69)	-	-	-	-
current age_1	6-11 months (88)	1.44	0.281	0.74	2.80
current age_2	12- 17 months (55)	1.16	0.700	0.54	2.45
current age_3	18+ months (150)	0.98	0.956	0.48	2.00
gender	Female (198)	-	-	-	-
gender	Male (164)	0.92	0.334	0.77	1.09
birth weight_0	1.2- <2.0 kg (16)	-	-	-	-
birth weight _1	2- <2.5 kg (36)	0.82	0.466	0.49	1.39
birth weight _2	2.5- <3.0 kg (109)	0.77	0.307	0.47	1.27
birth weight _3	3.0+ kg (197)	0.92	0.750	0.57	1.51
multiple babies_0	1 child (337)	-	-	-	-
multiple babies_1	Twins (22)	1.05	0.803	0.73	1.51
multiple babies_2	**Triplets (3)**	**1.77**	**0.028**	**1.06**	**2.93**
PCT*_agegroup_0	13-20 years (49)	-	-	-	-
PCT_agegroup_1	21-30 years (186)	1.14	0.379	0.85	1.51
PCT_agegroup_2	31-40 years (92)	0.98	0.896	0.71	1.35
PCT_agegroup_3	41-50 years (23)	0.96	0.851	0.62	1.48
PCT_agegrp_4	**51-63 years (12)**	**1.73**	**0.007**	**1.16**	**2.59**
education_0	no schooling (92)	-	-	-	-
education _1	1-7 years (156)	1.11	0.391	0.88	1.39
education _2	8-12 years (105)	0.96	0.715	0.76	1.21
education _3	>12 years (9)	0.61	0.067	0.36	1.04
literate PCT	No (130)	-	-	-	-
literate PCT	Yes (231)	1.00	0.991	0.83	1.21

Family assisted	No (273)	-	-	-	-
Family assisted	**Yes (89)**	**1.34**	**0.003**	**1.11**	**1.63**
breast feeding _0	0- 12 months (45)	-	-	-	-
breast feeding _2	13- 18 months (78)	1.13	0.436	0.83	1.54
breast feeding _3	19- 23 months (27)	1.01	0.946	0.68	1.51
breast feeding _4	24+ (212)	1.15	0.329	0.87	1.51
age at weaning _0	0-2 months (130)	-	-	-	-
age at weaning _1	3- 6 months (219)	0.85	0.107	0.71	1.03
age at weaning _2	7+ months (13)	1.33	0.164	0.89	2.00
Weight-age-z score_0	>-1 (214)	-	-	-	-
Weight-age-z score_1	-1.99 to -1 (92)	0.99	0.961	0.80	1.24
Weight-age-z score_2	≤ -2 (56)	1.01	0.962	0.79	1.29
Height-age-z score_0	>-1 (162)	-	-	-	-
Height-age-z score_1	-1.99 to -1 (97)	1.07	0.547	0.85	1.36
Height-age-z score_2	≤ -2 (103)	1.18	0.131	0.95	1.47
Weight-height-z score_0	>-1 (280)	-	-	-	-
Weight-height-z score_1	-1.99 to -1 (56)	0.82	0.141	0.63	1.07
Weight-height-z score_2	≤ -2 (26)	0.97	0.854	0.69	1.36
family_unit_0	2-5 members (122)	-	-	-	-
family_unit_1	6-7 members (77)	0.85	0.197	0.67	1.09
family_unit_2	8-10 members (80)	0.99	0.927	0.76	1.29
family_unit_3	11+ (83)	1.15	0.210	0.92	1.43
family_children_0	1-3 children (152)	-	-	-	-
family_children_1	4-5 children (105)	1.04	0.744	0.84	1.27
family_children_2	6+children (105)	1.09	0.437	0.87	1.36
siblings under 6 years_0	no siblings <6yrs (88)	-	-	-	-
siblings under 6 years_1	1-2 siblings<6yrs (212)	1.17	0.181	**0.93**	1.47
siblings under 6 years_2	**3-4 siblings<6yrs (53)**	**1.32**	**0.048**	1.00	**1.73**
siblings under 6 years_3	5+ siblings<6yrs (9)	1.34	0.178	0.87	2.06
siblings 6 years or more_0	1-3 siblings 6+ yrs (129)	-	-	-	-
siblings 6 years or more_1	4-6 siblings 6+ yrs (113)	0.84	0.144	0.68	1.06
siblings 6 years or more_2	7-10 siblings 6+ yrs (89)	1.11	0.380	0.88	1.40
siblings 6 years or more_3	11+ siblings 6+ yrs (31)	0.99	0.971	0.74	1.33

male_siblings_0	0-2 male siblings (260)	-	-	-	-
male_siblings_1	3-6 male siblings (95)	1.09	0.423	0.88	1.34
male_siblings_2	**7+ male siblings (7)**	**1.36**	**0.064**	**0.98**	**1.88**
female_siblings_0	1-4 female sibs (183)	-	-	-	-
female_siblings_1	5-9 female sibs (141)	1.07	0.452	0.89	1.30
female_siblings_2	10+ female sibs (38)	1.04	0.748	0.80	1.36
# of dead siblings	None (270)	-	-	-	-
# of dead siblings	1 (72)	1.09	0.439	0.88	1.34
# of dead siblings	2 (20)	0.86	0.478	0.56	1.31
child stays	mostly inside house (340)	-	-	-	-
child stays	mostly outside hse (22)	1.06	0.766	0.72	1.55
child's care_0	Mother (300)	-	-	-	-
child's care _1	another family member (41)	0.91	0.505	0.68	1.21
child's care _2	**House help (7)**	**0.45**	**0.008**	**0.24**	**0.81**
child's care _3	**School (1)**	**1.13**	**0.039**	**1.01**	**1.26**
child's care _4	mother / family member(13)	1.52	0.078	0.95	2.43
house type	block walled (116)	-	-	-	-
house type	**mud walled (246)**	**1.29**	**0.008**	**1.07**	**1.56**
house ownership_0	owner occupied (293)	-	-	-	-
house ownership_1	Rented (58)	0.83	0.145	0.64	1.07
house ownership_2	not owned/ rented (11)	0.94	0.766	0.62	1.42
toilet_type 0	no toilet (139)	-	-	-	-
toilet_type 1	**Flush (14)**	**0.52**	**0.002**	**0.34**	**0.78**
toilet_type_2	Latrine (209)	0.85	0.093	0.71	1.03
burn refuse	No (91)	-	-	-	-
burn refuse	yes (271)	0.99	0.945	0.82	1.21
main_fuel_0	gas/paraffin (11)	-	-	-	-
main_fuel_1	Charcoal (66)	1.07	0.824	0.61	1.88
main_fuel_2	Firewood (276)	1.37	0.239	0.81	2.33
main_fuel_3	firewood/charcoal (9)	1.26	0.590	0.53	2.97
cooking location_0	Outside (28)	-	-	-	-
cooking location_1	different house from sleep area (196)	0.96	0.812	0.69	1.33
cooking location_2	same house as you sleep (138)	1.01	0.961	0.72	1.42

water source_0	both well & piped (14)	-	-	-	-
water source_1	open/closed well (22)	1.37	0.246	0.80	2.34
water source_2	Piped (326)	1.23	0.384	0.77	1.96
watersite_0	own source (27)	-	-	-	-
watersite_1	Public (284)	1.19	0.309	0.85	1.66
watersite_2	Shared (51)	1.17	0.434	0.79	1.72
job description of MIP_0	non-skilled (145)	-	-	-	-
job description of MIP_1	Trade (74)	1.17	0.168	0.93	1.48
job description of MIP_2	Skilled (115)	0.85	0.147	0.68	1.06
job description of MIP_3	**Professional (28)**	**0.71**	**0.029**	**0.52**	**0.96**
smokers in HH_0	None (255)	-	-	-	-
smokers in HH_1	1 (83)	0.99	0.897	0.78	1.24
smokers in HH_2	**2 or more (24)**	**1.47**	**0.005**	**1.13**	**1.93**
adults in index's room_0	None (11)	-	-	-	-
adults in index's room_1	1 adult (97)	0.92	0.706	0.58	1.44
adults in index's room_2	2 adults (242)	0.94	0.793	0.62	1.45
adults in index's room_3	3+ adults (12)	0.81	0.481	0.44	1.47
# siblings <6yr living in same house as index_0	no siblings < 6 yrs (139)	-	-	-	-
# siblings <6yr living in same house as index_1	1 sibling (164)	1.07	0.512	0.87	1.30
# siblings <6yr living in same house as index_2	2-4 siblings (59)	1.01	0.941	0.78	1.31
# siblings 6+ years living in same house as index_0	no siblings 6+ yrs (134)	-	-	-	-
# siblings 6+ years living in same house as index_1	1-2 siblings (158)	1.02	0.882	0.84	1.25
# siblings 6+ years living in same house as index_2	3-4 siblings (57)	0.87	0.285	0.67	1.12
# siblings 6+ years living in same house as index_3	5-7 siblings (13)	0.98	0.867	0.65	1.43
# siblings <6yr sleeping in same room as index_0	no siblings < 6 yrs (190)	-	-	-	-
# siblings <6yr sleeping in same room as index_1	1 -3 sibling < 6 yrs (172)	1.08	0.373	0.91	1.30
# siblings 6+ years sleeping in same room as index_0	no siblings 6+ years (250)	-	-	-	-
# siblings 6+ years sleeping in same room as index_1	1 siblings (70)	0.96	0.776	0.74	1.25
# siblings 6+ years sleeping in same room as index_2	2-4 siblings (42)	0.81	0.117	0.62	1.05

# siblings <6yr sleeping in same bed as index_0	no siblings < 6 yrs (224)	-	-	-	-
# siblings <6yr sleeping in same bed as index_1	**1-3 siblings (138)**	**1.21**	**0.038**	**1.01**	**1.45**
# siblings 6+ years sleeping in same bed as index_0	no siblings < 6 yrs (309)	-	-	-	-
# siblings 6+ years sleeping in same bed as index_1	1-2 siblings (53)	1.18	0.240	0.89	1.56
# siblings <6yr going to school_0	no siblings < 6 yrs going to school (289)	-	-	-	-
# siblings <6yr going to school_1	1 sibling (53)	0.88	0.303	0.68	1.12
# siblings <6yr going to school_2	2-3 siblings (20)	1.40	0.059	0.99	2.00
# siblings 6+ years going to school_0	no siblings < 6 yrs going to school (109)	-	-	-	-
# siblings 6+ years going to school_1	1- 2 sibling (143)	0.97	0.749	0.78	1.19
# siblings 6+ years going to school_2	3-5 siblings (98)	0.97	0.828	0.76	1.25
# siblings 6+ years going to school_3	6-9 siblings (12)	1.17	0.482	0.76	1.80

*PCT (Primary care taker)

§p-values (bold<0.05) for the tests of the null hypothesis that each hazard ratio is one.

Appendix X: Testing the Proportional Hazard Assumption

Test and graphs based on Schoenfeld residuals

1. RSV infection

Time: Log(t)

Variables	rho	chi2	df	Prob>chi2
PCT age group _1	-0.01754	0.08	1	0.7748
PCT age group _2	-0.00454	0.01	1	0.9387
PCT age group _3	-0.00574	0.01	1	0.9242
PCT age group _4	-0.00092	0.00	1	0.9889
care giver _1	-0.03683	0.40	1	0.5246
care giver _2	0.08581	1.57	1	0.2107
care giver _3	0.04153	0.34	1	0.5583
care giver _ 4	-0.06639	1.01	1	0.3153
family assisted	0.04269	0.39	1	0.5310
house_toilet_1	-0.0437	0.33	1	0.5659
house_toilet_2	-0.02495	0.17	1	0.6793
house_toilet_3	-0.05702	0.94	1	0.3315
house_toilet_4	-0.01654	0.03	1	0.8692
house_toilet_5	-0.00969	0.02	1	0.8951
smokers in HH_1	0.10178	3.08	1	0.0791
smokers in HH_2	-0.05376	0.43	1	0.5099
global test		7.60	16	0.96

2. RSV-LRTI model

Time: Log(t)

Variables	rho	chi2	df	Prob>chi2
house_toilet_1	0.04997	0.36	1	0.551
house_toilet_2	-0.02583	0.08	1	0.7718
house_toilet_3	-0.05431	0.41	1	0.5203
house_toilet_4	0	.	1	.
house_toilet_5	0.02908	0.14	1	0.7078
care giver _1	-0.0303	0.1	1	0.7551
care giver _2	0.07026	0.45	1	0.5007
care giver _3	0	.	1	.
care giver _4	-0.08484	0.46	1	0.4994
family assisted	-0.02106	0.04	1	0.8326
smokers in HH_1	0.12797	2.33	1	0.1269
smokers in HH_2	-0.10636	1.23	1	0.2679
PCT age group _1	0.20378	6.21	1	0.0127
PCT age group _2	0.01011	0.01	1	0.9088
PCT age group _3	0.1041	1.35	1	0.2445
height for age z-score _1	0.00893	0.01	1	0.9178
height for age z-score _2	0.18827	3.58	1	0.0585
siblings under 6 years_1	-0.13606	1.92	1	0.1659
siblings under 6 years_2	-0.02614	0.11	1	0.7346
siblings under 6 years_3	-0.07343	0.45	1	0.5037
family_children_1	0.14783	2.98	1	0.0844
family_children_2	0.00562	0	1	0.9505
job description of MIP_1	0.0641	0.37	1	0.541
job description of MIP_2	0.10201	1.24	1	0.2657
job description of MIP_3	0.05525	0.42	1	0.5148
global test		21.71	25	0.6524

3. LRTI model

Time: Time

Variable	rho	chi2	df	Prob>chi2
current age_1	-0.01115	0.10	1	0.7473
current age_2	-0.00955	0.09	1	0.7620
current age_3	-0.00894	0.09	1	0.7656
education _1	0.00011	0.00	1	0.9966
education _2	0.00284	0.01	1	0.9120
education _3	-0.03226	1.18	1	0.2768
height-age-z score_1	0.00054	0.00	1	0.9865
height-age-z score_2	0.01837	0.37	1	0.5436
family_children_1	0.01198	0.21	1	0.6476
family_children_2	-0.06846	7.94	1	0.0048
main_fuel_1	-0.01687	0.33	1	0.5638
main_fuel_2	0.01571	0.29	1	0.5900
main_fuel_3	0.00106	0.00	1	0.9694
# siblings <6yr sleeping in same room as index_1	-0.03445	1.57	1	0.2099
# siblings <6yr sleeping in same room as index_2	0.00379	0.02	1	0.8942
# siblings 6+ years sleeping in same room as index_1	0.03362	1.60	1	0.2063
# siblings 6+ years sleeping in same room as index_2	0.07068	6.54	1	0.0106
# siblings 6+ years sleeping in same room as index_3	-0.00222	0.01	1	0.9357
# siblings <6yr going to school_1	0.03548	1.42	1	0.2338
# siblings <6yr going to school_2	-0.00073	0.00	1	0.9786
global test		29.24	20	0.0831

Appendix Y: Assessment of Goodness of fit of the Cox regression models.

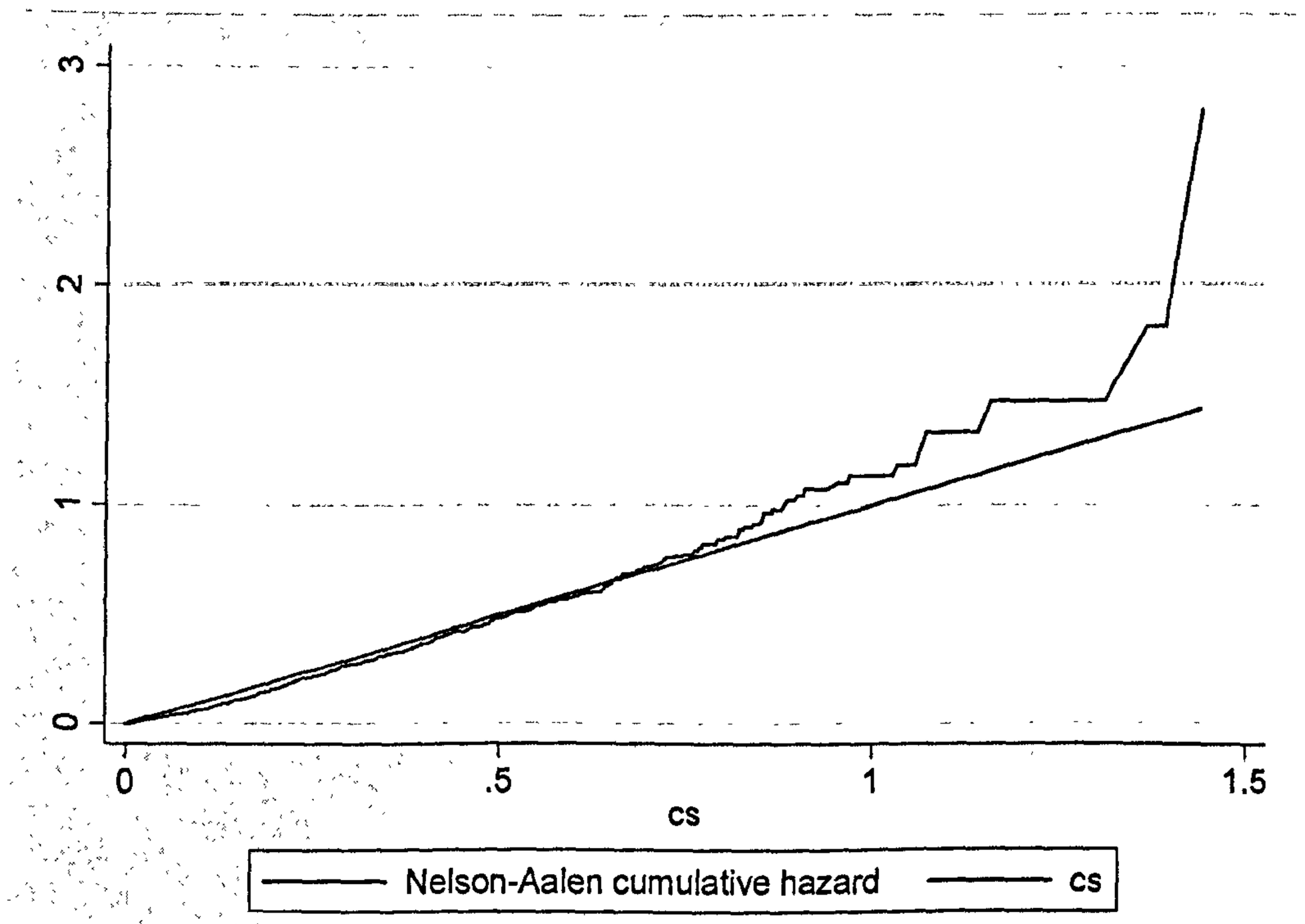


Figure 1: Plot of Nelson-Aalen cumulative hazard function and the cs residuals for RSVI

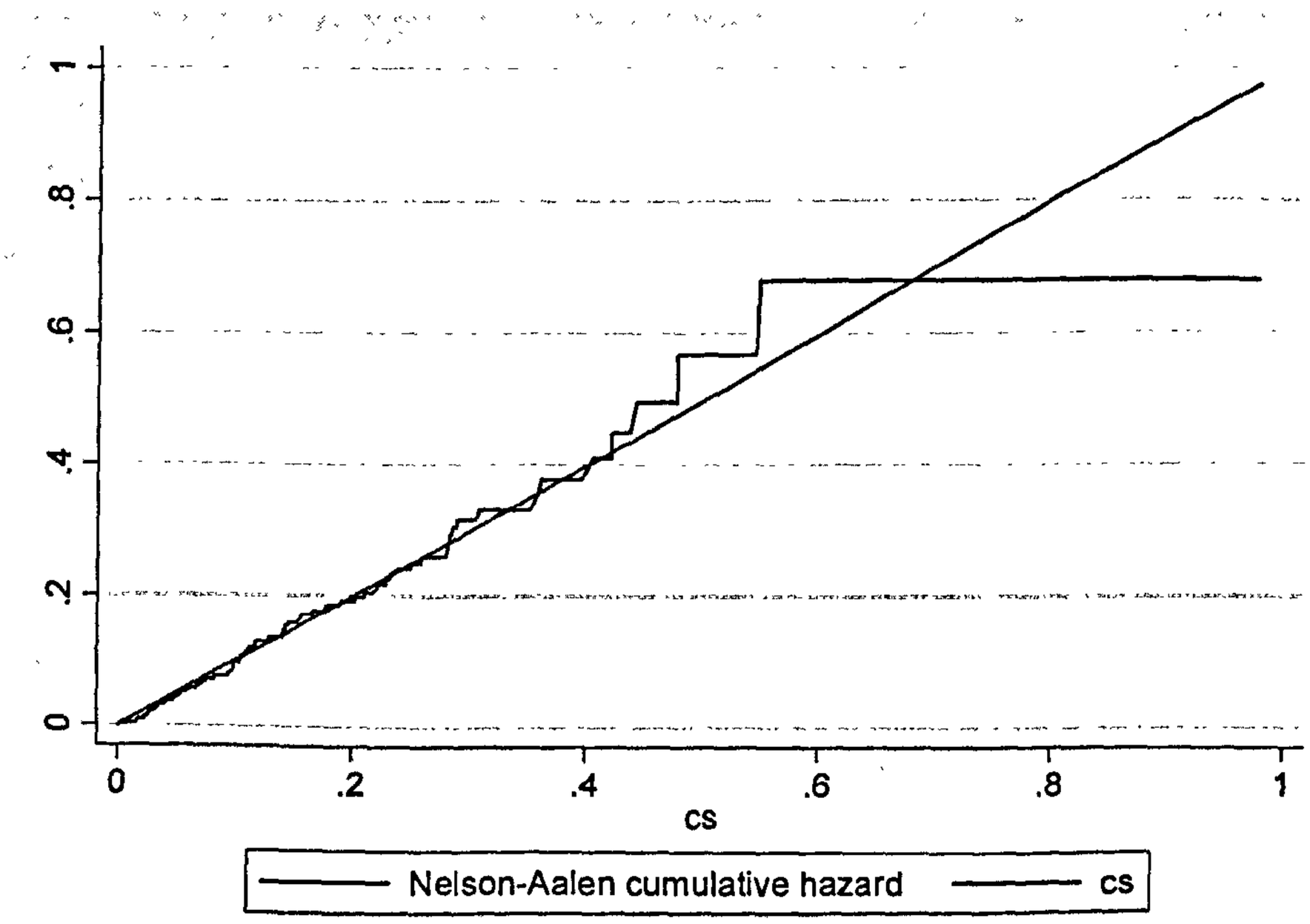


Figure 2: Plot of Nelson-Aalen cumulative hazard function and the cs residuals for RSV-LRTI

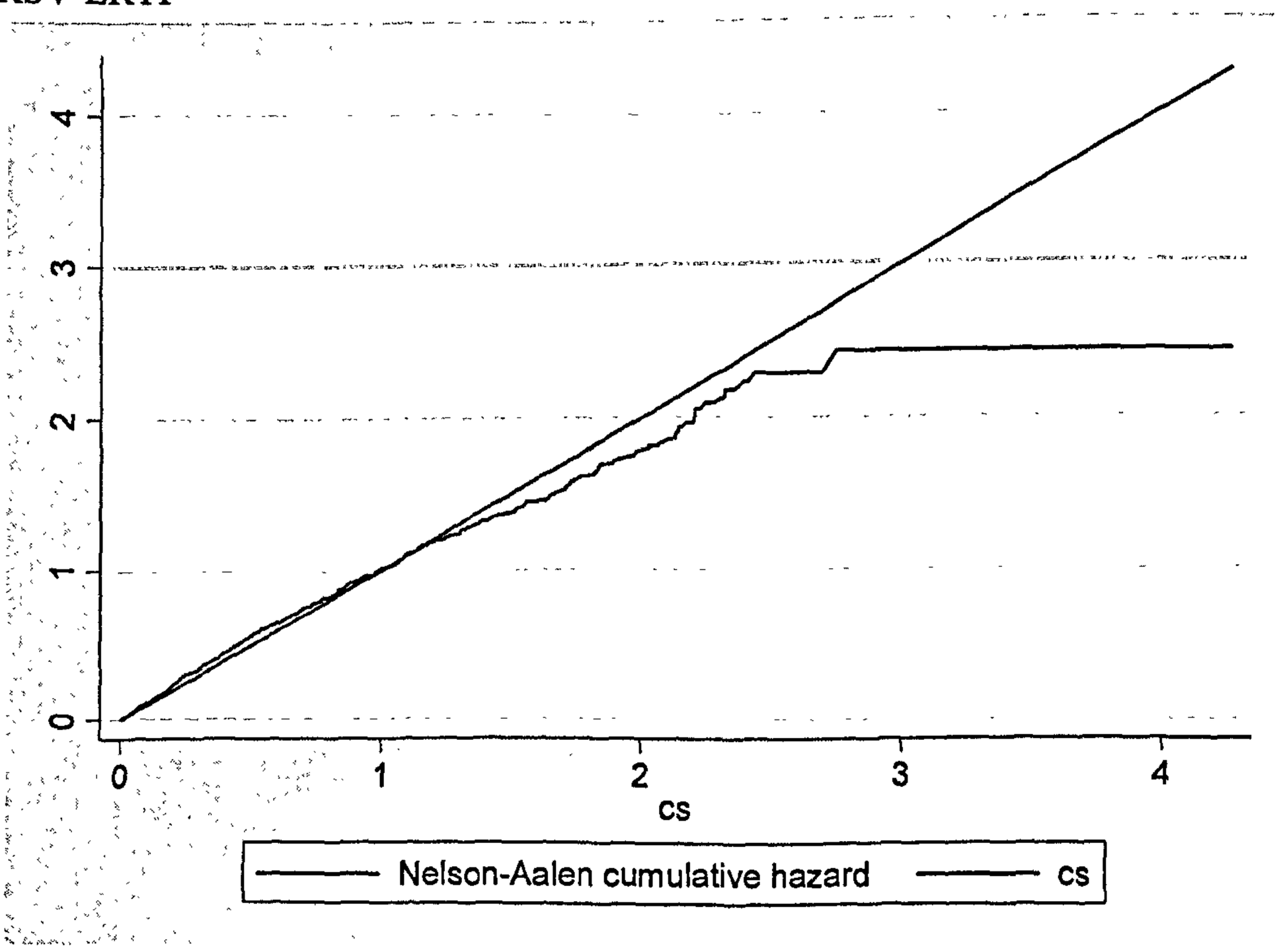


Figure 3: Plot of Nelson-Aalen cumulative hazard function and the cs residuals for LRTI